

North Korea's way with weapons

The deal between the United States and North Korea is not the climb-down it may appear, but should spur the nuclear powers to a comprehensive test-ban treaty before next year.

LAST week's agreement on nuclear weapons between the United States and North Korea is not the happiest agreement of its kind. The US administration's political opponents are muttering about the deal. The International Atomic Energy Agency (IAEA) in Vienna is uneasy at the prospect of the delay there will now be before it will have access to sites it wishes to inspect. There is even more general discontent that the example of North Korea will now become an illustration to the world's smallest powers of how to win important concessions from the United States: pretend to have a nuclear weapons programme, reject inspections under IAEA safeguards agreements and wait for an invitation to Geneva (where last week's deal was signed).

Luckily, that is not the whole truth, or even half of it. North Korea is a special case in at least three respects. First, it is an ex-client state of the ex-Soviet Union which is now out in the economic cold. (Economic production is actually falling, as in Cuba and other states in the same plight.) Second, North Korea is ethnically and culturally a part of the Korean peninsula, of which economically successful South Korea occupies the remainder. Third, the North Korean government is in the same fix as China some years ago: how can it follow the path of economic liberalism while preventing its own people from breaking the strict rules under which they live?

The concessions wrung from the United States last week, amounting to more open diplomatic recognition and the modification of the US trade embargo in force since the shooting stopped in the Korean War in 1953 (there is still no peace treaty), will not in themselves resolve those dilemmas. The most likely possibility is that South Korean companies will begin to invest in the north, but Pyongyang will be nervous that only the subversive dream of reunification can follow. And, of course, the suspicions are correct. Long before the two light-water reactors promised by the United States have been commissioned (against the promise not to extract plutonium from fuel already irradiated), it is more likely than not that North Korea will be a very different place.

Even so, the worry that the North Korean example will provoke other members of the Non-Proliferation Treaty (NPT) to follow suit is not unreal, although the most obvious analogies are inapplicable. Countries such as Cuba, for example, would not at this stage be able to launch even a credible rumour of nuclear intentions on the world. A more likely outcome is that, if the time should come for brokering an agreement between, say, India and Pakistan, they will

join the NPT, the brokers will be faced with demands they cannot meet.

The lessons, for the nuclear powers, are straightforward. The most urgent need is that they should be prepared to disarm the mostly empty grumbling by non-nuclear states about the asymmetry of the NPT by finishing off the negotiation of a comprehensive test-ban treaty. With the NPT review conference due next April, there is not now much time left. The second lesson is that the other declared nuclear powers (there are four of them) should wring out of China a much more explicit declaration of its strategy in respect of nuclear weapons. There was a nuclear test (underground) in China on 7 October, but nobody is saying why. □

Who is Caesar's wife?

Politicians accused of conflicts of interest may be a model for ethics in the research community.

As the research profession awakens to the potentially corrosive influence of external commercial interests within academic institutions, it must take some note of the charges of corruption now flying about in public life. Sadly, the election of a new government in Italy earlier this year has not banished long-standing anxiety about financial links between the government and those who seek to influence it. In Japan, the present government owes its place to the taint of influence-peddling its predecessor had acquired. In the United States, there is the usual trickle of fall-out from public office as people are alleged to have broken creditably strict rules. (Representative Dan Rostenkowski is the latest to be under a cloud.) Now the rot seems to have spread to France and Britain, where ministers have been compelled by the discovery of wrong-doing to give up their posts in recent weeks and days.

The cynical view that politicians are always and everywhere corrupt is a slander, but its wide currency is all too understandable and must, at least subliminally, help to shape the way that people elsewhere behave. But, superficially at least, a study of political behaviour would offer little help in guiding academic researchers through the conflicts of interest they will encounter. Politicians are usually caught out selling influence, sometimes for personal reward, often in their party's interest. (The British minister who quit last week was apparently paid £2,000 for asking a question in the



House of Commons.) A peculiarly British anxiety is that the government, having hived off much public administration to unelected committees, is now accused of filling them with people whose political friendship is more stringently tested than their competence.

But the common and offensive thread in this behaviour pattern is that people who are elected for a specific purpose, to represent their electors' interests, and who are paid a salary in recompense, end up representing some interests more vigorously than others, or arranging that some public duty should be discharged less than competently. These are simple breaches of the implied contract people running for election offer voters when seeking office. It does not matter that representatives are free to vote on specific issues according to their best judgement and their consciences. To be required to make an honest and unbiased judgement can also be a public responsibility. In that sense, the accusations against politicians are no different from an employee's breach of fiduciary responsibility to an employer. Britain exaggerates its own difficulties by allowing Members of Parliament to moonlight, often for parliamentary lobbyists.

How do those transgressions compare with growing indifference to conflicts of interest arising in research? To the extent that politicians' misdemeanours are a betrayal of public trust, they are in a class of their own. But the easy acceptance of conflict in research can also mean betrayal. This journal's puritanical view that academic researchers should never be officers of a separate commercial company, even one they set up themselves, has logic on its side; they would have to neglect their regular duties if the company were in trouble. But what of those who become stockholders in return for membership of an advisory board, or who favour one of several suitor companies with a consultancy contract?

The common practice, at least in universities, is that the institution will agree provided that regular duties are not neglected, which is usually enforced by limiting the time spent on outside work. But what if a person earns a substantial fraction of his or her annual salary in ways like that? Or if the stock appreciates remarkably? Such developments can readily undermine the trust on which members of productive laboratories mutually rely for their success. Will colleagues be free with their half-baked bright ideas if they do not know which company will hear of them next? In the research enterprise, there can be no alternative to the full disclosure of outside commitments, not only to institutional authorities but also to working colleagues.

But would not all academics then migrate to industry? That reflects many British politicians' defence of moonlighting in the House of Commons; salaries are so poor that members have to supplement their incomes. The simple solution is that they should be paid adequately, and that legitimate extra earnings (from speaking engagements and the like) should then be strictly limited, as they are in the United States. Why not follow the same recipe in research? A cap on consultancy and other earnings would be only seemly. Decent pay is long overdue if the research profession is not always to be overshadowed by those with longer pockets. □

Another rope trick

The University of Chandigarh in India has been disgracefully lenient with Professor V.J. Gupta.

PROFESSOR V.J. Gupta is now a famous person, not on the strength of his many spurious contributions to the literature of the palaeontology of the Himalayas, but because of the length of time during which the Panjab University of Chandigarh has failed to face the Gupta problem. Five years have passed since the case against Gupta appeared in print (see J.A. Talent *Nature* **338**, 613–615; 1989). Now, after a two-year inquiry, Judge M.S. Gujral says that he has found "overwhelming" evidence that much of Gupta's published research is indeed fraudulent. And the punishment? The university senate has reprimanded Gupta, denied him further salary increments and decreed that he shall not in future hold an administrative post in the university. (By seniority, Gupta had been the next to take a turn as dean of the university.)

This is not simply a way of heaping one scandal on another; it is also a source of shame for the whole research community in India. It is particularly shocking that a person found to have put false research data into circulation should be allowed to continue teaching students, research students included. The university senate cannot avail itself of the excuse that Gupta would have been able to fight a more severe penalty through the tortuous Indian courts. Had it been externally constrained, it would not gratuitously have decreed that four of Gupta's co-authors should be stripped of whatever credit they may have earned in that unfortunate position.

It will be interesting and important to see what happens next. The best outcome would be that Gupta should simply resign, but that is simply an expression of the hope that the problem will go away. The university senate could also follow last month's timorous decision by calling for such a course of action. That, at least, would prevent Gupta from claiming that the mild penalties now decreed must be a sign that he has been innocent all along. But even that seems unlikely. Only five of the 55 members of the senate voted for a motion for his dismissal when the question came up last month. Indeed, it seems improbable that there will be a solution within Chandigarh.

That is why external influences should be mustered. In these unprecedented circumstances, it is not unreasonable that academic institutions elsewhere, even outside India, should make their opinions known to the university whenever they have dealings with it. Cutting off funds from the centre, on the other hand, would harm students at an institution that serves an valuable purpose regionally and (in agriculture) nationally. Meanwhile, *Nature* will keep in touch with Gupta's co-authors, much persecuted in the past five years, and will protest in public if they are further humiliated. At some stage, it must be hoped, the university at Chandigarh will begin to worry about its reputation. □

Boston University claims scientific successes built on earmarked funds

Boston. As November's mid-term elections draw close, the US Congress and its activities have become subject to an unprecedented barrage of public criticism. But the beleaguered institution has found itself a friend at Boston University.

In choosing sites for major new scientific facilities, Congress is "the only appropriate body to decide how federal tax dollars should be spent," says John Silber, the university's president. The alternative — peer review by a panel of scientists — "has conflict of interest built into it almost all the time", he says.

Silber is a fiery iconoclast who stood as Democratic candidate for governor of Massachusetts in 1990. His views are not universally shared; M. R. C. Greenwood, a senior official in the Office of Science and Technology Policy, warned a recent congressional hearing that such views "have the potential of destroying the process" that underpins science in the United States.

But looking across the Charles River from Silber's offices, the appeal of earmarking to this middle-ranking but politically well-connected school is all too obvious. On the opposite bank sits the Massachusetts Institute of Technology (MIT), and half a mile up the road is Harvard University, both well-endowed icons of academic excellence, and both well placed to succeed in any process based on peer review.

In recent testimony to Congress, Silber branded these elite schools "an informal cartel," and compared them with the oil producers' organization OPEC. Membership of the elite, he argued, has been built on first-rate facilities such as MIT's Lincoln Laboratory — funds for which were earmarked by Congress in 1952.

"Earmarking is the way all of the major research universities got their start," says Silber. Now it's Boston University's turn: since 1985, the school has attracted three awards of earmarked funds for the construction of research facilities, worth a total of \$56.5 million.

The first, for a science and engineering centre, has made possible the rapid expansion of the university's engineering school. The second, for a physics and biology building, has helped raise the physics department from a position of relative obscurity to one of acknowledged strength. The third and largest is for \$29 million to help build a major photonics research centre that could make Boston University the leader in a critical new field of applied science.

In each case, the earmarked funds have enabled the university to attract a larger amount of loans to build the facility. The grant for the photonics centre, for example,

includes \$16 million towards total construction costs of almost \$80 million. "The federal government is going to get research facilities that cost four or five times what they've put in", Silber says.

Don Fraser, a former engineer at MIT and senior official at the Department of Defense (DoD), who was appointed last year to head the new photonics centre, says that it will emphasize the applications of photonics research. These, he says, go far beyond fibre optic telecommunications to embrace a range of exciting possibilities in medicine, environmental control and consumer electronics.

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Silber (right) praises direct funding of his university's new science and engineering centre (above)

The process by which the centre won funds from the Department of Defense was complex. After extensive lobbying by Boston University — much of it, Silber says, at the offices of Senator Edward Kennedy (Democrat, Massachusetts) — the \$29 million was included in the defence budget bill passed in 1991.

The Bush administration later sought (unsuccessfully) to rescind funding for the centre among 23 defence earmarks, worth \$170 million. Reviewing the photonics project in 1992, a DoD review panel rejected it: according to Representative Martin Hoke (Republican, Ohio), the panel's unpublished report branded the proposal as "superficial and not unique".

But at the beginning of 1993 the new DoD team, led by Les Aspin, decided to take a softer line on earmarks. DoD then helped all 23 projects to re-submit proposals, which the panel reviewed and accepted.

"I'm not going to sit here defending that process — it's a superb process," says Silber. "If someone comes up and asks what's two plus two and gets five, and comes back six months later and says 'I made a mistake, two plus two is four', then it is not to be criticized, it is to be complimented."

Silber also defends his use of a Washington-based lobbyist, Gerald Cassidy, who receives \$400,000 a year from the university. Cassidy's role, he says, is to find out where in the government money might be available, and to arrange meetings for Silber and his staff with the right people. Silber says that the actual selling of the projects is down to himself and his academic staff.

Past earmarking at Boston is now bearing fruit. The amount of peer-reviewed research attracted by the College of Engineering, for example, has grown from \$1.4 million in 1985 to \$9.2 million this year.

An even greater expansion has occurred in the physics department, although Silber's recent claim to Congress that it "is now ranked in the top six in the nation for the quality and influence of its research" oversells the progress made under Larry Sulak, who took charge of the department in 1985.

Silber's claim is based on an assessment by the Institute of Scientific Information (ISI) in Philadelphia. This ranked Boston University seventh in the number of citations per published paper across all the physical sciences (not just physics) in the period 1987–90 (see *Nature* 349, 6; 1991). The main physics earmark was granted by Congress in 1988.

"If we didn't have this building, we couldn't have done anything," says Sulak of the \$42 million physics and biology centre, which the \$8-million earmark helped to build. According to more recent ISI figures for the period 1981–93, Boston's physics department had the seventeenth highest citation rate among large physics departments in the US. The number of papers grew rapidly through the period, while their citation rate remained stable.

But the advantages of such a centre are plain for all to see. They include a machine shop, an electronics design and construction facility, and a computer centre, each superbly equipped. There is no money in the main federal science and technology budget for such university-based facilities, and has not been for twenty five years.

In a ringing endorsement of Kennedy, who is standing for re-election, in the Boston Globe on 11 October, Silber vigorously attacked his Republican opponent, Mitt Romney, and praised the veteran senator's clout in Washington, listing expensive federal projects he had brought back home. Oddly, the \$56.5 million Kennedy has brought to Boston University was absent from the list. "Romney calls this pork," Silber wrote. "I call it the return from Washington of our taxes." **Colin Macilwain**

Boston University Photo Services

Boston University Photo Services

Warning of retaliation over access to physics facilities

London. European countries that restrict the access of researchers from Japan and the United States to nuclear physics facilities built with European funds risk retaliation against their own scientists, according to the report of a group of advisers to the Organization for Economic Cooperation and Development (OECD).

The warning has been triggered by a recent decision by the Institut Laue-Langevin (ILL) in Grenoble, France, to restrict access to scientific groups from member states, or groups collaborating with them. The decision was made by the ILL's steering committee in the light of tight budgetary constraints and increasing demand for access to its beam lines.

The warning of retaliation also coincides with a meeting next week of representatives from OECD member states at the Oak Ridge National Laboratory (ORNL) in Tennessee. The meeting will pursue recommendations on future neutron sources made at an expert meeting on synchrotron radiation sources and neutron beams organized by the OECD's Megascience Forum in Risø, Denmark, late last year.

Topics to be addressed at next week's meeting, at which the report of the Risø meeting will be published, include closer collaboration on the planning and design of advanced neutron facilities, instrumentation for existing beamlines and access to current and future neutron sources.

Two main machines are under consideration. The United States is drawing up plans for a powerful research reactor, the \$2.7-billion Advanced Neutron Source (ANS), at ORNL. At the same time, Europe is considering building a spallation (pulsed) neutron source, to be known as the European Spallation Source, whose location is still under discussion.

Many physicists feel that, given the costs of the two facilities — as well as the possible precedent set by ILL's recent decision — some agreement on how they might be 'shared' by scientists on both sides of the Atlantic needs to be negotiated in advance.

In addition to the ILL, the newly opened European Synchrotron Radiation Facility (ESRF) at Grenoble, also operates under rules that break with the traditional practice of providing open access to international facilities based purely on the scientific merit of research proposals.

The new rules at both ILL and ESRF are more restrictive than those for large facilities in Japan and the United States, sparking fears that these countries might retaliate. John Axe, for example, scientific programme director of the High Flux Beam Reactor (HFBR) at the Brookhaven National Labo-

ratory (BNL) in Long Island, New York, compares the situation to the imposition of import taxes by one country on goods made elsewhere.

Retaliation can soon escalate, says Axe, with countries being denied access to potential markets for their products. "We understand the financial pressures that have driven that decision [at ILL], but we are sorry to see it. I'm hoping that other facilities won't follow their lead and that ILL will eventually reconsider their position."

According to Peter Tindemans, chairman of the OECD Megascience Forum and director of research and science policy at the ministry of education, culture and science in the Netherlands, one of the major aims of the Oak Ridge meeting is to identify "ways out of this tension" between governments demanding a fair return on their investment, and others who want open access for all.

Bill Appleton of ORNL, associate director for the ANS, says that the issues of collaboration, coordination of instrumentation and access to facilities have to be resolved. "Given the way the costs of major facilities are going in this country, there won't be a spallation source built in the United States, so we'll need to share facilities", he says.

The International Union of Pure and Applied Physics (IUPAP) is also keen to promote close coordination through open access to facilities. Axe plans to present a statement to the Oak Ridge meeting on behalf of the US liaison committee of IUPAP, emphasizing the importance of this issue. It has been drafted in conjunction with US high-energy physicist, some of whom are concerned about possible restrictions on their access to the proposed Large Hadron Collider at the European Laboratory for Particle Physics (CERN) in Geneva. **Maggie Verrall**

Brussels suggests 'variable geometry' research projects

London. The European Commission (EC) in Brussels last week suggested to the 12 member states of the European Union (EU) that joint research programmes involving some — but not all — members should for the first time be included within the commission's five-year 'framework' programmes of collaborative research.

At the same time, the commission suggested that, in order to avoid individual member states from feeling left out of arrangements that are referred to in France as 'variable geometry', a balanced portfolio of multilateral programmes should be approved simultaneously by the commission.

Both suggestions were made by the commissioner for research, Antonio Ruberti. They are among a number of new developments being proposed by Ruberti and his staff to fulfil a mandate in the Maastricht Treaty to improve the co-ordination of research policies between member states.

Some member states themselves are wary of the implications of these moves, concerned that they could lead to an excessive concentration of power in the hands of the commission, and could also encourage a 'two-speed' Europe in research. Participants at a meeting of the commission's own science policy advisory committee (CREST) are said to have given the proposals a cool response at a meeting last Friday.

But the commission itself feels that the Maastricht Treaty has given it the green light to introduce measures designed to ensure an increasing convergence between the research policies of EU member states.

Areas in which Brussels officials feel that the potential for convergence exists include employment practices, policies on funding for large facilities, and the use of forecasting techniques in identifying research priorities. □

Cresson tipped for EU research post

Paris. **Edith Cresson (right), the former French socialist prime minister, is being tipped to succeed Antonio Ruberti as the European Union (EU)'s research commissioner when he steps down at the end of the year.**

According to unfirmed reports, Cresson and Jacques Santer, the new president of the commission, agreed on her nomination at a meeting in Paris last week. Cresson is said to have preferred the post of relations with Eastern Europe,

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

but to have lacked the necessary political support from Edouard Balladur, the French prime minister.

During her spell as prime minister, Cresson was a strong advocate of industrial policy. Her appointment could therefore lead to closer relations between the directorates for science (DGXII) and industry (DGIII), as well as a shift in EU research funding towards more focused social and industrial goals. D. B.

Concern grows over 'trade' in Russian fossils

Munich. A scientifically important fossil stolen more than two years ago from the Museum of the Palaeontological Institute in Moscow has been returned, thanks to the efforts of a German museum curator who identified it and fought for its return.

But around 50 others are still missing. And, with thefts occurring at other institutes as well, there is evidence that some researchers, faced with declining salaries, may be involved in the illegal export of fossil specimens from scientific collections.

The Palaeontological Institute, whose museum became the main repository for the former Soviet Union's fossil treasures, is part of the Russian Academy of Sciences. Its world-renowned collection of fossils was started by Peter the Great in the eighteenth century and includes important specimens from the Permian and Triassic periods.

Following the fall of communism in 1990 and the country's ensuing economic collapse, the institute has been seeking alternative sources of finance to make up for the inadequacy of its state funding. At the same time, individual scientists have been under pressure to seek new sources of income.

The institute was given the role of protecting Russia's palaeontological heritage. A committee of its scientists, headed by the deputy director, Igor Novikov, is responsible for examining all requests for export licences for fossils. If it approves a request, this is formally endorsed by the ministry for the protection of the environment and natural resources.

There is growing evidence, however, that this system is being circumvented. In one theft in spring 1992, for example, fifteen 240-million-year-old Triassic skulls of labyrinthodonts, now-extinct amphibian ancestors, were stolen from the institute's amphibian collection.

Some of the stolen specimens are holotypes — reference samples for their species — and of great scientific value. Some specimens also have a high market value. The locked showcase containing the stolen fossils had not been broken. "Someone from the institute was certainly involved in the theft", says Michael Shishkin, head of

the institute's amphibian division.

After Shishkin had reported the theft of the skulls in *Lethaia Forum*, a specialist journal, Rupert Wild, curator of the State Museum for Natural History in Stuttgart, Germany, became suspicious about a specimen bought by one of his volunteer technicians. Wild became suspicious that a *Thoosuchus jacoblevi* skull, bought for DM1,600 (US\$1,070) from a German fossil dealer, Joachim Woerdemann, might be one of the stolen items.

The identification number on the skull, which had been partially scratched out but was still visible under special light, confirmed that it had once belonged to the institute's collection.

Nevertheless, the institute took little action to secure the return of the missing fossils. "Both Shishkin and myself were abroad for extended periods, and no-one else was interested", says Novikov.

As a result, most of the effort to return the *Thoosuchus* skull came from Wild, who tried to have Woerdemann prosecuted.

Woerdemann, who has not returned telephone calls from *Nature*, has said that he has never bought any fossils from the institute. But he has admitted that the *Thoosuchus* skull obtained by Wild, which he claims to have bought from a private collector, may have originated from the museum.

No charges were brought against Woerdemann after the police said they did not have enough evidence to support a prosecution. But Woerdemann eventually agreed to buy back the skull, saying that he wanted to maintain good relations with the institute. It was collected in Germany last month by Novikov, who took it back to Moscow.

Scientists in the West, as well as many in

the institute itself, remain alarmed at how scientifically important fossils, including vertebrate fossils for which no export licences have been recommended by the institute's committee, are still leaving the country.

But there is also concern about the implications of the commercial links that the institute has explored in its efforts to raise

funds. In the spring of 1992, for example, the institute signed a cooperative agreement with a company called Kaminy Svetok ("stone flower"), to collect, identify and preserve fossils. The agreement specified that, in return, the institute would receive 20 per cent from all sales made through the company and that it would also receive 80 per cent of any sales of fossils from its own collection.

Some worry that involvement in a commercial project may put off potential collaborators. Others fear



A stolen Labyrinthodont skull, now safely back in Moscow.

this arrangement represents a potential conflict of interest that could seriously undermine scientific research by tempting the institute to sell off important items in its scientific collections.

Even greater concern has been generated by evidence that some members of the institute may have set up a purely commercial operation, known as the Applied Palaeontological Institute (API), as a vehicle for the sale of fossils. Export papers for a specimen of a Pareiasaur, an ancestor of the turtle, on sale for £40,000 (US\$65,000) through a dealer in the United Kingdom, bear this name, and carry the same address as the academy's institute.

When contacted by a potential buyer, a staff member at the academy's institute initially acknowledged the API's existence. But other officials later denied all knowledge of the API. "I would very much like to know what the truth is", says Shishkin.

The *Thoosuchus jacoblevi* skull is the only specimen so far to have been located and returned. Others are believed to be in the United States, Japan, the United Kingdom and Germany. A joint British-Russian committee has been set up to locate and return the rest of the missing collection.

Meanwhile, Novikov says that in the past month the institute has forbidden its staff to take on commercial commitments and has cancelled its agreements with Kaminy Svetok.

Toni Feder & Alison Abbott

\$1.8bn nuclear ignition facility for Livermore

Washington. The US Department of Energy (DoE) has approved the design stage of the proposed \$1.8-billion National Ignition Facility (NIF) at the Lawrence Livermore laboratory, securing the future of the California as a nuclear weapons research complex.

The decision to proceed with the \$1.8-billion facility, which will use lasers to induce self-sustaining fusion, comes after months of pressure from the Department of Defense on Hazel O'Leary, the energy

secretary, to proceed with the project in order to demonstrate her commitment to the nuclear weapons programme.

The DoE says that the NIF will be useful in astrophysics, in developing fusion power and in the safe maintenance of the nuclear stockpile. But critics say it will damage US efforts to halt the proliferation of nuclear weapons and that its chief function is to keep nuclear physicists in work pending any resumption of nuclear weapons development. C. M.

Germany may back military satellite project

Paris. The prospect of an independent European military satellite observation system moved a step nearer last week with reports that Germany is poised to back military space endeavours for the first time with a programme whose cost is estimated at between DM2 billion and DM10 billion (US\$1.3–\$6.69 billion) over the next decade. The United Kingdom is also said to be prepared to support a European programme.

Until now, France has led European efforts in military space. Its military space budget, the fastest growing such budget in the world, has soared from FF697 million (US\$136 million) in 1987 to more than FF4 billion this year. France already has a FF9.23-billion Syracuse military communications payload aboard France Telecom's Telecom-2 satellite. Early next year it will launch its first high-resolution spy satellite, the FF8-billion Helios-1 satellite.

Edouard Balladur, the French prime minister, recently reaffirmed that military space was a national priority. Indeed, the ministry of defence was last year given shared responsibility for the French space agency CNES. This institutionalized already strong links — both share the same industrial base, for example, while Helios-1 uses the platform developed by CNES for the SPOT-4 Earth observation satellite.

French enthusiasm for military observation satellites has a relatively simple explanation. France has nuclear weapons, and one of the functions of Helios is to act as a targeting system. Although the United Kingdom, Europe's only other nuclear power, has similar needs, it has so far relied on its close relationship with the United States for access to satellite data.

Until now, Germany has (like Britain)

turned down invitations to join France's military efforts in space. One reason is that Helios-1 can operate only in clear skies, whereas Germany's security interests generally lie beneath cloudy skies, for example in Eastern Europe.

Germany's shift in position has been prompted partly by the decision of its constitutional court to authorize the operation of German peacekeeping forces outside the area of the North Atlantic Treaty Organization (NATO), as it needs an observation capacity to support such forces. On the industrial front, Germany is keen to support its sizeable satellite industry, while technologically the infrared capacity of Helios-2 — scheduled for launch in 2002 — allows it to see through clouds.

Germany's shift also coincides with a growing interest in the Western European Union (WEU), Europe's joint security organization, in establishing a satellite spying capacity independent of the United States. In 1991, WEU committed ECU38.25 million over three years for the construction of a WEU satellite ground station at Torrejon, near Madrid in Spain, to verify arms agreements and to monitor both security crises and environmental changes.

At present, the centre's main activity is the processing of images — for example from Bosnia — obtained from the SPOT series of satellites and other commercial suppliers. But it will soon handle classified images from Helios-1 and become a 'mini Central Intelligence Agency' (CIA) for processing satellite images.

The centre has given European countries joint capabilities they could not have developed separately and access to images that the United States "would never have given

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

France's Helios-1 satellite is winning more European supporters

them", says Stephane Chenard of the space consultancy Euroconsult in Paris.

France also needs European collaboration to afford its costly military programmes. It has paid 80 per cent of the costs of Helios-1 (Italy contributes 14 per cent and Spain 6 per cent) but cannot afford the FF10-billion Helios-2 alone, much less its planned FF11-billion Osiris radar satellite, scheduled for launch in 2004.

Moreover, Spain last week pulled out of Helios-2 for budgetary reasons, leaving the French with a gap in funding, while Italy, which is reviewing its defence policy, has yet to decide whether to participate.

But this gap now seems likely to be filled by Germany, which is expected to announce its participation in Helios-2 at a Franco-German summit in November.

Officials from the relevant ministries in Germany are refusing to comment publicly on their plans. But newspaper reports suggest that the participation is linked to an agreement that Germany should lead the Osiris programme, in which it will make a financial investment of up to 50 per cent.

The possible rapprochement between German and French efforts in military space is also indicated by a proposed merger between the satellite and missile divisions of Aerospatiale of Paris and Deutsche Aerospace, which will probably lead to the creation of a joint subsidiary in Munich.

As so often in European affairs, the joker in the pack is the United Kingdom, which has privileged access to CIA satellite data. But Britain apparently now recognizes that it shares security interests with France and seems ready to invest in Helios-2.

In principle, Britain does not seem opposed to a European military space effort. The industrial advantages are obvious, while military bodies, faced with strong budget pressures, are searching for a pragmatic, rather than a political solution, say observers. Increasingly, Europe wants its own missile warning system in order to reduce dependency on US goodwill in sharing data, they say.

Declan Butler

ESA agrees to extra year for ERS-1

Paris. The European Space Agency (ESA) has agreed to extend the operation of its remote-sensing satellite ERS-1 to the end of 1995 instead of shutting it down at the end of this year. The decision follows protests from scientists that shutting down ERS-1 in December would throw away an opportunity to carry out unique research, at little cost, using both satellites.

Researchers had argued that flying ERS-1 and ERS-2 — which is scheduled for launch in January next year — 'in tandem' would, for example, allow interferometric studies to be carried out on large swathes of the globe, including all of Antarctica (see *Nature* 370, 317; 1994).

ESA officials say that the agency has agreed to spend another ECU2 million to keep ERS-1 operational for another year in a 'safe orbit'. But member countries are still divided over whether to pay a further ECU6

million for the operational costs of a dedicated 'tandem' mission.

ESA says this sum could easily be found by taking funds from ERS-2 and reducing its lifetime by a few weeks, although some political problems remain because the countries paying for ERS-1 are not identical to those paying for ERS-2.

But, despite caution from several ESA members over the financial aspects, the German space agency DARA says that the "scientific and technological case [for a tandem mission] has now been made", and is optimistic that the extra money will be found.

ESA officials also point out that even without extra funding for the operational costs of a dedicated tandem mission, the new lease of life given to ERS-1 means that it will in fact now function in tandem with ERS-2 all next year.

D.B.

UK/Japan projects in life sciences to target priority areas

London. Britain's Office of Science and Technology (OST) has been asked by the government to identify high priority areas for collaborative projects in the biological sciences that could be carried out jointly with research groups in Japan.

The proposal to build closer links between the two countries in the life sciences is one of several recommendations in a report published last week by the OST's newly-formed UK/Japan and Asia-Pacific Advisory Group.

Announcing his support for many of the group's conclusions, David Hunt, the minister for science, said that while existing links had brought substantial benefits to both countries, further attention now needed to be paid to "ensuring that collaborative efforts are in priority areas."

Hunt said that it would be to the advantage of both countries if the areas for collaboration were deliberately targeted. "In this way we can ensure that resources are directed as positively as possible."

Sir William Stewart, the government's chief scientific adviser, said that this approach is similar to that taken by the government last year in giving a focus on wealth creation to the research councils, in particular the new Biotechnology and Biological Sciences Research Council (BBSRC). "This is a continuation into the international arena of what we have done with the BBSRC," says Stewart.

At the same time Denis Noble, professor of physiology at the University of Oxford, and a member of the advisory panel which draw up the report, pointed out that there was a need to maintain links with Japan "outside the targeted areas".

The task of helping to identify areas for potential collaboration has been passed to Sir John Cadogan, the director-general of the research councils, although the group recommends that the final selection of key areas be left to the research councils.

Other recommendations included in the report are that existing schemes to support the exchanges of scientists and technologists should be extended to cover a wider range of disciplines, including engineering; and that a group of British and Japanese patent experts carry out a joint study of patent policies in the two countries, exploring ways in which the two systems might be brought closer in line.

The report also confirmed the British government's view that it would like Japan and the United Kingdom to enjoy "equal status" as members of international efforts such as the European Laboratory for Particle Physics (CERN) and of future neutron and spallation sources. **David Dickson**

Japan's universities 'need to strengthen links to industry'

Hikone, Japan. In contrast to the situation in the United States, research in Japanese universities is still isolated from the outside world and does not respond sufficiently to social needs, according to a leading Japanese industrialist.

These views were expressed by Michiyuki Uenohara, executive adviser of NEC Corporation, at a symposium held last week in Hikone, central Japan, and attended by about two dozen university presidents from the two countries and representatives of Japanese industry.

In a panel discussion on interactions with industry, Uenohara said that university researchers in Japan should increase their interaction with the outside world and explore basic research from "the real world".

But Junjiro Kanamori, president of Osaka University, reminded participants that Japan's national universities are formally "branch organizations" of the central government. As a result, he said, the government always "casts a shadow" on relations between universities and industry.

Although links between Japanese universities and industry have grown slightly over the past decade, thanks largely to several new initiatives by the Ministry of Education, Science and Culture, most of the funds for these programmes — ¥52 billion (US\$690 million) out of ¥67 billion in fiscal year 1994 — consist of untied 'donations' from industry. These are primarily used by companies as a tool for recruiting good students and seldom encourage interaction.

According to Uenohara, there are some fundamental flaws in the research environment in Japanese universities that discourage interaction with industry. For example, interdisciplinary research is "rarely" found in the universities. Yet advances in the information technology that Uenohara's company produces depends on research bridging not only natural science and engineering, but also medical and social science.

Kanamori echoed this view, emphasizing that universities must first "get over the sectionalism within themselves" before interactions can flourish. Hiroo Imura, presi-

dent of Kyoto University, also pointed out the lack of mobility among Japanese university researchers compared with those in the United States. Furthermore, he said, Japanese companies "do not employ PhDs", resulting in little cross-fertilization between university and company researchers.

Government regulation also inhibits collaboration. Uenohara pointed out that the present interactions are largely "unidirectional". Industrial researchers can carry out research on campuses, but university researchers are prohibited by law from participating in research projects in industrial laboratories.

Furthermore, they can carry out consultancy work for only one or two hours a week as they are civil servants with government pay cheques. As a result, very few university faculty members understand industrial research or can explain their research in industrial terms.

Several US participants expressed envy of the long-term view of research that Japanese industrial leaders such as Uenohara can take. Japanese industry has in recent years been able to increase its investment in basic research because of the low interest rates in Japan during the bubble economy of the late 1980s. Furthermore, unlike those in the United States, Japanese companies do not have to answer to their stockholders in the short term.

In contrast to the United States, however, government financial regulations severely inhibit venture businesses and entrepreneurship in Japan, according to Naoya Yoda, executive adviser to Toray Corporate Business Research Inc. In particular, restrictions on the flow of funds into venture businesses holds back interaction between universities and industry.

But probably the biggest factor preventing such interaction is university researchers themselves. In reply to a question after the meeting, Kanamori said he does not get the impression that university researchers are pushing for regulations that prevent them from working in company laboratories to be eased. **David Swinbanks**

Caskey to head Merck's genome efforts

Washington. Tom Caskey, the geneticist who has played a leading role in recent moves to establish a map of the human genome available for 'unencumbered' use by researchers in both the private and the public sectors, is to join the US pharmaceutical company Merck as senior vice president for basic research.

Caskey is leaving the department of

molecular and human genetics at Baylor College of Medicine in Houston, Texas, to take charge of all Merck's basic research programmes, as well as the human genome sequencing effort that the company is financing at the Washington University in St Louis, Missouri, announced at the end of last month (see *Nature* 371, 365; 1994). **C. M.**

Tunisian institute to tackle secrets of malaria genome

Washington. Daniel Cohen, head of the team that produced the first-generation physical map of the human genome and director of the Centre des Etudes du Polymorphisme Humain (CEPH) in Paris, is stepping up a campaign to establish a genomics research institute in Tunisia.

The institute would work on the genome of *Plasmodium falciparum*, the bacterium responsible for the most severe form of human malaria which kills between three and four million people annually.

Cohen's aim is for the Tunisian institute to be the first of several in developing countries that promote technology transfer through high quality research on diseases of importance to the developing world. Part of the plan is to train and employ native-born scientists.

Jeffrey Ravetch, head of the Laboratory of Biochemical Genetics at the Sloan Kettering Institute in New York, would be responsible for the institute's scientific programme. Ravetch has been involved in malaria research for 10 years and over the past three years has applied techniques developed by the human genome project to the *P. falciparum* genome. He has written the science proposal for Ifriqya, a fund-raising foundation set up by Cohen two months ago in France and named after the ancient name for Tunisia, which he intends to establish soon in the United States.

Ravetch's laboratory is one of five in the United States, the United Kingdom and Australia that are working on *P. falciparum* with funding from the Burroughs-Wellcome Trust. He emphasizes that the science proposal for Ifriqya is still at a preliminary stage. But he adds: "If all goes ahead, the Tunisian institute's work would be rolled into our current effort."

Cohen, who is making a major commitment to the Tunisian venture, will next week seek financial backing from Burroughs-Wellcome. He is also planning to approach wealthy French industrial families and is organizing fund-raising concerts at which he plays the piano.

Since first formulating his ideas for high-tech transfer two years ago, Cohen has raised \$1 million of the \$5 million he needs to build the institute, and the Tunisian government has donated the land.

Cohen's plans received a setback when representatives of the French government on CEPH's board recently turned down his request to put one or two per cent of CEPH's budget into the project. But Cohen says that the French research minister, François Fillon, has told him that at CEPH's next meeting this veto will be withdrawn.

Ravetch says that the Ifriqya initiative has four broad scientific aims: to develop

the expression tags for cloned-DNA libraries of the organism's genes (perhaps 3,500 genes) and create a physical map; to develop a genetic map; to develop techniques for transforming the parasite during the only

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Malaria: still major killer.

stage of its life cycle that can currently be cultured *in vitro*; and finally to sequence the whole genome.

Cohen is now recruiting members to Ifriqya's scientific board and hopes to attract well-known figures such as Jacques Cousteau and James Watson, who he says have given him much intellectual encouragement.

Helen Gavaghan

French minister bends to protests over CNRS funds

Paris. François Fillon, the French minister of research, last week announced that he was making available FF200 million (US\$39 million) of emergency funds to alleviate the financial crisis at the Centre National de la Recherche Scientifique (CNRS).

He is also trying to find additional money to eliminate the need for CNRS's controversial decision that researchers reduce their spending to 60 per cent of the funds they had been allocated for the year. The decision had led to widespread protests throughout France (see *Nature* 371, 639; 1994).

At the same time, Fillon said that "dismantling" CNRS is "out of the question", although he does not rule out change. This response was prompted by concern over proposals contained in an internal government circular to shift control of half of CNRS's 1,300 laboratories to universities.

The circular had also suggested switching half the funds that CNRS laboratories now receive to a series of new programmes awarded to groups on the basis of peer review of grant proposals.

But Fillon said that concern about reforms of CNRS is "premature" as these are still "non-existent".

D. B.

Stanford ends research costs row

San Francisco. The US Office of Naval Research (ONR) has dropped charges of fraud and wrongdoing against Stanford University in California after the university agreed last week to pay the US government \$1.2 million to settle a four-year-old dispute over its billing of indirect costs for government research between 1981 and 1992.

The university had already paid \$2 million to the government shortly after ONR accused it in 1990 of overcharging by up to \$185 million during the previous decade. It admitted to 'errors' that include depreciation for a yacht that had been donated to the student sailing programme, part of a scandal that led to the resignation of Stanford's then president, Donald Kennedy.

University officials said last week that they felt relieved and vindicated by the low level of the final settlement accepted by the government, and the fact that the charges against it had been dropped. "With [the settlement] behind us, Stanford can devote its attention fully to its ongoing mission of teaching and research on the frontiers of knowledge", Gerhard Casper, the university's president, said in a statement.

The government has now set in place a series of regulations addressing the ambi-

guities that underlie the disagreement over Stanford's billing system. "The parameters are much clearer", said Lieutenant Commander Kenneth Ross, a spokesman for the Navy. Stanford had not acted with fraudulent or criminal intent, he said, "but it was not the way we perceived things should be done".

The ONR negotiates with 39 US universities on indirect cost billing rates on research for itself and all other federal agencies. The money covers the costs of administration, libraries, student services and any other item that cannot be directly charged to a particular grant or contract.

Under a new rule announced last week, ONR said it would shift to a multi-year approach in its negotiations. This is the last of more than 17 changes in the rules that the government has introduced since the Stanford dispute highlighted the issue.

Reed Brimhall, director of the office of government costs and rate studies at Stanford, says that the new rules have generally worked to shift costs from the government to universities. For example, he says, a new limit on administrative expenses and student services is unrealistic and is exceeded by most major research universities in the United States.

Sally Lehman

Study proposed on integrity of published data

Washington. A member of the recently created US Commission on Research Integrity last week suggested that a confidential study be carried out to check the correspondence between raw data and those appearing in published papers, in order to determine the extent of scientific misconduct.

The idea of such a cross-check was first put forward in 1989 by Desmond Rennie, professor of medicine at the University of California, San Francisco, and a commission member. He made the proposal in an article in the *Journal of the American Medical Association*, of which he is a deputy editor.

At a meeting of the commission in Washington DC, Kristina Gunsalus, vice-chancellor for research at the University of Illinois at Urbana-Champaign and another member, said that she was prepared to draw up the new proposal in order to "take the heat off Rennie", acknowledging that his idea had been "met with howls of protest".

The idea of such a study was supported by Kenneth Ryan, chairman of the commission, whose creation was mandated by Congress in last year's National Institutes of Health (NIH) Reauthorization Act. "If everyone was told when they submitted a paper that their data might be requested confidentially, I would hope that it would be obvious that they would agree," Ryan said later.

Gunsalus claims that a study of the integrity of published data would not be intended as an attack on scientists. But she points out that there is no hard data to support the scientific establishment's frequent assertion that there is little misconduct in science.

Congress' decision to require NIH to set up the commission was prompted by frustration with the way in which allegations of scientific misconduct have been treated and concern that the NIH's Office of Research Integrity (ORI), which investigates allegations of misconduct, has not been working as well as Congress had hoped.

Ryan's panel will investigate the competence of the ORI and advise on definitions of misconduct, and on how allegations should be treated and whistleblowers protected. It will report to Donna Shalala, the Secretary of Health and Human Services, who has responsibility for the NIH.

Ryan, who is a professor of obstetrics, gynaecology and reproductive biology at Harvard Medical School, believes that the whistleblower issue will be tough. "Some of the things done to whistleblowers, and sometimes by distinguished scientists, are unconscionable," he says. "We know of cases where people are fired. They complain to the NIH, which insists they are reinstated, but it is a hollow victory because they are reinstated in a closet down the hall."

The commission met for the third time last week and is at the beginning of a two-year process, but it has so far attracted little attention. "It has been treated as a non-event, with a view of 'been there, done that'. I think this is because people don't know who Ken Ryan is," says Gunsalus.

Ryan is, in fact, an old hand at sorting out thorny ethical issues, and has encountered both scepticism from the media and Congress, and defensiveness from the biomedical research community. For example, he is a former chairman of the National Commission for the Protection of Human Subjects in Biomedical and Behavioral Research. "We were viewed then as either whitewashers, or as revolutionaries who would kill science," says Ryan.

Misconduct is currently defined as fabrication, falsification, plagiarism and other practices. It is the last phrase that causes most contention, as it covers loosely issues such as authorship, the degree of supervision a mentor should provide for a graduate student, how long data is kept and where.

The current commission is already leaning towards a two-tier approach where federal intervention would be reserved for egregious cases and the scientific community would accept responsibility for establishing common standards of behaviour across disciplines and institutions. **Helen Gavaghan**

NASA scientists brace for new cuts

Washington. Faced with a squeeze on funding for basic research, the managers of the US National Aeronautics and Space Administration (NASA)'s planetary programme last week took the unusual step of calling in a panel of outside scientists to informally advise them on where and how the research budget might be cut.

At least for the next year or two, the group was able to suggest temporary solutions that do not require cutting into the peer-reviewed 'core programmes' on which a large proportion of the US planetary science community relies for grant money. But the scientists who attended the meeting say that a larger problem still exists, namely the fact that basic research is raided whenever the space agency finds itself in a fiscal jam.

The budget for planetary research and analysis (R&A) is around \$100 million a year, relatively little by NASA standards. But it provides grants for hundreds of researchers around the United States working on data from spacecraft missions like Voyager and Magellan. It also supports activities as diverse as instrument development and looking after the lunar samples brought back by the Apollo astronauts.

It is these non-research activities that have caused much of the problem. NASA has recently committed itself to several new projects that fall outside the traditional R&A domain. These include participation in the Keck Observatory (see *Nature* 371, 189; 1994) and supplying an instrument to fly on Japan's Planet-B Mars mission in 1998. But with no extra funds allocated for the new

projects, they have been imposed on an R&A account that is already strained.

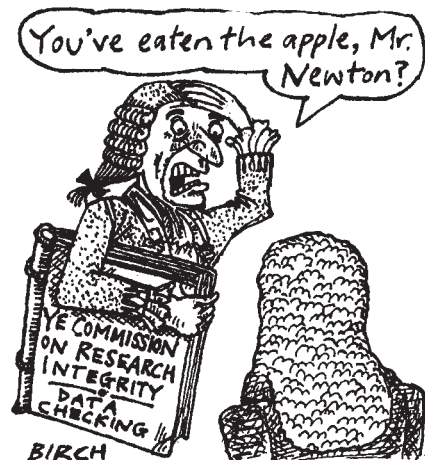
Faced with the prospect of cutting research grants to accommodate the new spending, NASA R&A managers asked for outside advice. After three days of deliberation, the panel, chaired by Jonathan Lunine of the University of Arizona, agreed on several possible solutions, which will be put to NASA next month.

For example, rather than introducing new research efforts (each with its own funding requirements) to address results from the Clementine lunar mission and the Shoemaker-Levy impact with Jupiter, the group proposes incorporating these subjects into existing research programmes. It also agreed to recommend that technology development within the R&A programme be streamlined and consolidated.

These two proposals, it calculates, will be sufficient to avoid the near-term budget shortfalls. But beyond such accounting games, the scientists also had to come to terms with the new realities of NASA. Belt-tightening is currently taking place across the whole agency, including high-profile programmes such as the space shuttle.

Joseph Burns of Cornell University, who chairs the National Research Council's Committee on Planetary and Lunar Exploration, says planetary research is especially vulnerable to reductions in NASA funding because there are so few tenured academic positions in the field. "A lot of people are on soft money," he says. "Everybody's in pretty fragile shape."

Tony Reichardt



ESF works, don't change it

SIR — The European Science Foundation (ESF), based in Strasbourg, is a non-governmental collaboration between the major national research funding organizations in Europe. ESF has 55 member organizations from 20 European countries, a staff of 29 and an overall budget for 1994 of FFfr73.5 million. Of this budget, FFfr37.9 million goes directly to promoting research, in the form of programmes and networks. Part of the rest is spent on developing new proposals.

We became in 1992 the co-chairmen of the programme REHE (Relativistic Effects in Heavy-Element Chemistry and Physics, planned for 1993–97), grouping 12 countries and having a yearly budget of about FFfr 0.8 million. This sum buys 2–5 workshops a year, contributes to a Euro-conference every second year and made possible 26 short-term visits in 1993. Most of these visits led to one or more publications that otherwise would not exist. That output/input ratio seems to us extremely efficient.

The allocations are made by a steering group or, empowered by them, the co-chairman. Most correspondence is by e-mail or fax. If necessary, a decision can be made in a day and the money actually paid out by Strasbourg in 2–3 weeks.

We know of few other organizations on this medium-sized scale and with this efficiency and speed. Obviously one requires a subject that still has unsaturated collaboration potential, and personal chemistry that works.

We have recently learned that the activities of ESF are being discussed, the issue being an eventual larger role in science policy matters. It is not within our competence to say whether this is desirable, but we would definitely regard it unwise to do so at the cost of ESF's concrete scientific activity.

Pekka Pyykkö

Department of Chemistry,
University of Helsinki,
PO Box 19 (Et. Hesperiankatu 4),
FIN-00014 Helsinki, Finland

Bernd Hess

Department of Physical and Theoretical
Chemistry,
University of Bonn,
D-53115 Bonn, Germany

No need to scoff

SIR — As a foreigner who has spent three and a half years working at a British university, I feel reasonably acquainted with British idiosyncrasies. Yet I was taken aback by the vehemence of your reaction to the plans for an expatriate

university in Thailand, and especially by the references to culture (*Nature* 371, 462; 1994). The article, it occurs to me, would be more fitting in an issue of *Nature* of a hundred years ago. Culture is being exported in many forms all across the world, not least by "international journals". If the plan is properly managed, the possibilities offered by a university could compensate for many such exports of lesser value. The fact that the United States has universities abroad should not automatically cause traditionalist British scientists to scoff at the idea.

Werner Sieber
rue du Botzet 3,
1700 Fribourg,
Switzerland

SIR — The export of British universities to faraway lands is not a new phenomenon — many former British colonies have well-established universities that began as offshoots of British universities. Most have survived and adapted, with a gradual replacement of expatriate staff by well-qualified local academics. Research does pose a dilemma, because in a new university there is a temptation to emulate illustrious (and expensive) facilities overseas only to find that the reality of short-term expatriate appointments prevents useful research from being completed. But robust and useful research is possible, particularly when it takes into account the context in which the new university is set — a new set of research possibilities will be perceived by open minds. So perhaps the founding of a new university in Thailand is not a time for cynicism, but an opportunity to nurture a valuable new resource.

Richard Dryden

Tor & South West College of Health,
Kinninmont Centre,
West of England Eye Infirmary,
Magdalen Street,
Exeter EX2 4HT, UK

Divergent proteins and views

SIR — Alexander Rich¹ says that Linus Pauling, together with Emile Zuckerkandl, in 1962, "by looking at changes in the amino-acid sequence of haemoglobin in different animals, could correlate the changes with the evolutionary period at which the animals diverged from each other", so as to measure evolutionary time. "It was a startling idea when it was presented."

Actually, the idea was old and well-established. The first molecular comparisons of haemoglobins were made in 1907 and 1909 by E.T. Reichert and A.P. Brown^{2,3}. They found, for example, that cats, dogs, a bear and a seal could be

placed in four divergent groups on the basis of axial ratio in crystals of their reduced haemoglobins. Their findings were widely quoted in textbooks.

In 1961, V.M. Ingram⁴ published a phylogenetic tree showing the successive separation of myoglobin and alpha, gamma, beta and delta haemoglobins from a common ancestor.

The concept of "molecular disease" was introduced by A.P. Garrod's publication on inborn errors of metabolism in 1908. It was Hardin Jones at the University of California, Berkeley, who first pointed out the dangers of low-level radiation from nuclear testing; indeed, Pauling lists a reference to Jones⁵ in his book⁶.

Thomas H. Jukes

University of California,
Berkeley,
California 94720, USA

1. Rich, A. *Nature* 371, 285 (1994).
2. Reichert, E.T. & Brown, A.P. *Proc. Soc. exp. Biol. Med.* 5, 66–75 (1907).
3. Reichert, E.T. & Brown, A.P. Monograph, Carnegie Institute of Washington, Washington DC (1909).
4. Ingram, V.M. *Nature* 189, 704–709 (1961).
5. Jones, H.B. University of California Radiation Laboratory Publication No. 3105 (1955).
6. Pauling, L. *How to Live Longer and Feel Better* (Freeman, New York, 1986).

Experimenting on the vulnerable

SIR — A recent issue of *Monthly Nature* contained an article suggesting male contraception by vasectomy and cryopreservation of sperm for eventual artificial insemination¹. But the first steps in implementing this otherwise attractive scheme, as proposed by the authors, are that "... the military services should begin a large-scale sperm cryopreservation programme".

I think that we should remember that, as agreed in international guidelines on human experimentation², military personnel belong to a vulnerable group. Such people (prisoners, medical students, children, those with mental disorders) should not be used for medical research because of their restricted freedom to make independent choices.

Despite the authors' view¹ that army personnel have a higher risk of early death and may therefore be interested in preservation of sperm, I do not think that this is a good argument for approving experimentation on soldiers and on prisoners.

Vasily Vlassov

(Colonel of Medical Corps)
Committee on Biomedical Ethics,
Saratov Medical University,
PO Box 1528,
Saratov 410601,
Russia

1. Djerassi, C. & Leibo, S.P. *Nature* 370, 11–12 (1994).
2. International Ethical Guidelines for Biomedical Research Involving Human Subjects (CIOMS, Geneva, 1993).

DNA fingerprinting dispute laid to rest

Eric S. Lander and Bruce Budowle

Two principals in the once-raging debate over forensic DNA typing conclude that the scientific issues have all been resolved.

THE US public, usually indifferent to matters scientific, has suddenly become obsessed with DNA. Nightly newscasts routinely refer to the polymerase chain reaction (PCR) and even the tabloids offer commentary on restriction fragment length polymorphisms (RFLPs). The new-found fascination with nucleic acids does not stem from recent breakthroughs in genetic screening for breast cancer susceptibility or progress in gene therapy — developments which will indeed affect the lives of millions. Rather, it focuses on the murder case against the former US football star, O. J. Simpson.

The Los Angeles trial, starting in November and to be broadcast live by several major television networks, will probably feature the most detailed course in molecular genetics ever taught to the US people. This bold experiment in public education should, in principle, be a cause for rejoicing among scientists. The catch is that the syllabus is being prepared by attorneys whose primary roles are as adversaries; the likely result is confusion. Already, the news weeklies are preparing the ground with warnings that DNA fingerprinting remains "controversial", being plagued by major unresolved scientific issues.

Forensic DNA typing certainly did provoke controversy soon after it was introduced into US courts in 1988. The technology itself represents perhaps the greatest advance in forensic science since the development of ordinary fingerprints in 1892, and is soundly rooted in molecular biology. The problem, however, stemmed from the manner of its introduction. Pioneered by biotech start-up companies with good intentions but no track record in forensic science, DNA typing was marred by several early cases involving poorly defined procedures and interpretation¹. Standards were lacking for such crucial issues as: declaring a match between patterns; interpreting artefacts on gels; choosing probes; assembling databases; and computing genotype frequencies. There is broad agreement today that many of these early practices were unacceptable, and some indefensible. For its part, the US Federal Bureau of Investigation (FBI) moved much more deliberately in developing procedures, sought public comment and opted for conservative procedures.

As a result of these growing pains, forensic DNA typing was subjected to

intense debate and scrutiny. When it first burst on the scene, the supporting scientific literature consisted of a mere handful of papers. By the middle of this year, there had been more than 400 scientific papers, 100 scientific conferences, 3 sets of guidelines from the Technical Working Group on DNA Analysis Methods (TWGDAM), 150 court decisions and, importantly, a 3-year study by a National Research Council (NRC) committee released in 1992 (ref. 2). In the light of this extraordinary scrutiny, it seems appropriate to ask whether there remains any important unresolved issue about DNA typing, or whether it is time to declare the great DNA fingerprinting controversy over.

As co-authors, we can address these questions in an even-handed manner. B.B. was one of the principal architects of the FBI's DNA typing programme, whereas E.S.L. was an early and vigorous critic of the lack of scientific standards, and served on the NRC committee. In a world of soundbites, we are often pegged as, respectively, a "proponent" and an "opponent" of DNA typing. Such labels greatly oversimplify matters, but it is fair to say that we represent the range of scientific debate.

We recently discussed the current state of DNA typing, and could identify no remaining problem that should prevent the full use of DNA evidence in any court. What controversy existed seems to have been fully resolved by the NRC report, the TWGDAM guidelines and the extensive scientific literature. The DNA finger-

printing wars are over.

Our goal is to correct the lingering impression to the contrary. Our analysis below represents our unanimous opinions (apart from specific comments about the workings and intent of the NRC committee itself, which necessarily are based on E.S.L.'s recollection). We focus on the subject most often said to remain problematical: population genetics. Our main thesis is that the academic debate that continues to swirl about population genetic issues is rooted in a misunderstanding of the NRC report and is, in any case, of no practical consequence to the courts. We also touch on how the legal and scientific community should cope with the continuing evolution of DNA typing technology. In particular, we question whether a steady succession of *ad hoc* committees, however distinguished, is a wise solution.

Laboratory practices

The initial outcry over DNA typing standards concerned laboratory problems: poorly defined rules for declaring a match; experiments without controls; contaminated probes and samples; and sloppy interpretation of autoradiograms¹. Although there is no evidence that these technical failings resulted in any wrongful convictions, the lack of standards seemed to be a recipe for trouble. To address these problems, the NRC committee enunciated conservative standards for each laboratory step, based on more than a decade of experience with human DNA



Comparing autoradiographs from DNA samples at Cellmark Diagnostics.

analysis. TWGDAM also developed guidelines along similar lines. Today, there is no doubt about the correct laboratory protocols to ensure reliable DNA typing results. Since the NRC report, US courts have unanimously accepted the technical reliability of DNA evidence, both in principle and in practice.

The NRC committee also highlighted the importance of laboratory accreditation, rigorous quality assurance and quality control (QA/QC) programs, and external blind proficiency tests (tests administered by persons outside the testing lab itself). The importance of these practices has been universally acknowledged, and most forensic labs follow TWGDAM's voluntary quality-control guidelines.

Population genetics

The controversy over population genetics began as a secondary issue. If DNA analysis reveals that two samples match at the loci tested, the final step is to estimate the frequency of the shared genotype in the general population, which indicates the probability that a randomly chosen person would carry this genotype. Such estimates depend on surveys of the appropriate population.

In some early cases, the rarity of genotype frequencies was greatly overstated owing to a technical error: the calculations were based on overoptimistic assumptions about the precision with which genotypes could be measured. One commercial lab, for example, reported the astronomical frequency of 1 in 738,000,000,000,000, based on a four-locus match¹. The NRC committee easily rectified these problems by requiring consistency between the measurement precision used for forensic analysis and for population genetic estimates (a practice that the FBI, in fact, had long followed).

A subtler but more challenging issue emerged in later cases, concerning the structure of human populations. The 'product rule', used by forensic labs to calculate genotype frequencies, assumed that the individual alleles comprising a genotype could be treated as statistically independent, and their frequencies multiplied². However, some population geneticists asserted that the assumption of independence was appropriate for well-mixed populations (technically, those at Hardy-Weinberg equilibrium and linkage equilibrium), but was not necessarily valid for populations with substructure. According to this argument, the frequency of a common Japanese genotype might be underestimated because the product rule ignored the fact that common Japanese alleles tend to occur together in the US Asian population. Moreover, the frequency of genotypes arising from mixed ethnic ancestry might be understated because the product rule was typically applied to separate racial databases

(for the Caucasian, Black and Hispanic populations) and thus did not account for the presence of genotypes involving common alleles from different racial groups. The substructure argument became a *cause célèbre*, pitting such luminaries as Lewontin and Hartl³ against Chakraborty and Kidd⁴. Both sides conceded that substructure could matter in principle, but many doubted that its effect could be significant in practice (see ref. 5).

The NRC committee at first attempted to settle the issue on its merits. The members agreed that the product rule was probably near the mark, but were hard pressed to say just how close. The committee considered applying formulas from theoretical population genetics based on the empirical measures of the degree of variation and admixture among and within populations. However, it concluded that there were, at the time, too few hard data about the loci used in forensic typing (most classical genetic surveys concerned protein polymorphisms, likely to be strongly influenced by natural selection) and about the precise structure of the US population. It would be too risky to base a recommendation on assumptions that might subsequently turn out to be faulty.

Thomas Caskey (Baylor College) eventually pointed the way out the quagmire when he asked, out of frustration, whether it was possible to ignore population substructure altogether. Taking up the notion, the NRC committee set out to fashion an extremely conservative rule having the virtue that it made virtually no assumptions.

The ceiling principle

The solution turned out to be quite simple. Suppose that the US population is descended from a collection of populations $P_1, P_2, \dots, P_n$, each sufficiently old and well mixed to allow the product rule to be safely applied. Regardless of the population substructure, the multiplication rule requires only a slight modification to yield a strict upper bound on the frequency of any genotype G : for each allele in G , the allele frequency should be taken to be the maximum over the component subpopulations. In effect, the approach makes the worst-case assumption that the population may contain individuals who, for example, carry a common Caucasian allele at a locus on chromosome 2 and a common Black allele at a locus on chromosome 17. By assuming the worst, one is guaranteed to be conservative. Because it used the maximum frequency in any subpopulation, the method was dubbed the 'ceiling principle'².

In practice, it is unnecessary to survey every possible subpopulation. The committee concluded that the likely variation in allele frequencies could be reckoned by conducting modest surveys of 100 individuals from each of 10–15 representative

subpopulations spanning the range of ethnic groups represented in the United States — such as English, Germans, Italians, Russians, Navahos, Puerto Ricans, Chinese, Japanese, Vietnamese and west Africans. Each allele frequency could then be taken to be the maximum over these subpopulations, although never less than 5%. (The latter provision was designed to deal with unexamined populations. If an allele was rare in the 10–15 subpopulations surveyed, genetic drift was not likely to have caused its frequency to drift much above 5% in other significant subpopulations.) Even in advance of detailed data about ethnic groups, the committee felt that same principle could be applied to the available racial databases (Caucasian, Black, Hispanic, Asian), although it recommended a 10% floor on allele frequencies to reflect the greater uncertainty about subpopulation variation: this slightly amended form was called the 'interim ceiling principle'. (The choices of 5% and 10% were based on the quantitative effect of genetic drift on the match odds — that is, on the reciprocal of the allele frequency — although none of this reasoning survived into the text of the final report.) The practical effect of these rules was to limit the contribution of any single locus to a factor of 50:1 odds based only on aggregate data for racial classifications and 200:1 odds based on more detailed ethnic surveys.

The ceiling principle was unabashedly conservative. It gave the benefit of every conceivable doubt to the defendant, so that it could withstand attacks from the most stubborn and creative attorneys. Some of the statistical power was sacrificed to neutralize all possible worries about population substructure.

The committee was comfortable with such a lop-sided approach, because even these extreme assumptions did not undermine the practical use of DNA fingerprinting. A four-locus match performed by forensic labs could still provide odds of 6,250,000:1. If this were not enough, two additional loci could increase the odds to more than 15,000,000,000:1.

Finally, the ceiling principle was not intended to be exclusive. Expert witnesses were still free to provide their statistical "best estimate" of genotype frequencies based on the product rule. But if disagreement over such estimates arose, the ceiling principle provided an approach that all parties had to admit was biased to favour a defendant. By all rights, this seemingly solomonic solution should have ended the controversy over population genetics.

Hitting the ceiling

Surprisingly, attacks came from an unexpected quarter. Some vocal theoretical population geneticists and statisticians concluded that the committee had been too conservative. They argued that the

effect of population substructure was slight and that it would best be treated by using formulas from theoretical population genetics. The ceiling principle was accused of being clumsy and scientifically flawed. Suddenly, a new controversy over population genetics seemed to emerge⁵⁻¹⁰.

Peter Menze/SPL

The debate was based on a simple misunderstanding of the NRC Committee report but, with the committee disbanded, there was no easy way to address it. Moreover, the committee members had agreed to let the report speak for itself to avoid the emergence of conflicting gospels according to different members. In retrospect, this was probably an unwise decision because it has allowed a minor academic debate to snowball to the point that it threatens to undermine the use of DNA fingerprinting by suggesting that there is some problem with the use of population-genetic estimates in court.

Six objections have been raised to the ceiling principle, which are worth briefly refuting:

(1) **The ceiling principle is premised on the flawed analysis of Lewontin and Hartl that there is significant population substructure⁸.** On the contrary, the committee was quite dubious that substructure had significant effects, but felt that the possibility needed to be taken seriously rather than dismissed based on theoretical or indirect arguments. The NRC report cites the Lewontin-Hartl article³ only twice, both times balanced against a longer list of opposing articles.

(2) **The ceiling principle is scientifically flawed because it is not used in population genetics^{9,10}.** Moreover, the plan to sample 10-15 representative populations is statistically unsound⁸. The choice of a statistical method necessarily depends on the dangers of overestimation versus underestimation. In forensics, there is strong agreement on the need to be conservative. By contrast, population geneticists do not need to be conservative in academic studies; they are content to err equally often on the high and low side, and thus ceiling approaches are unnecessary. However, ceiling approaches are common in a closely related genetic pursuit: the mapping of disease genes. To guard against falsely implicating or excluding a chromosomal region, human geneticists often analyse their data under worst-case conditions, such as using an unrealistically low ceiling on the penetrance of a disease gene or an unrealistically high frequency for a marker allele^{11,12}. Also, some authors have complained that surveys of 100 individuals do not allow accurate estimates of the frequency of rare alleles⁸. The purpose of the suggested population studies, however, was not to estimate low frequencies, but rather to check that some alleles do not unexpectedly have extremely high frequencies (much more than 5-10%) in certain populations. For this purpose, 100

individuals is quite adequate.

(3) **The ceiling principle makes ludicrous assumptions about the possible substructure of a population^{9,10}.** Although the NRC report called for empirical studies of those groups that made significant contributions to the United States (such as English, Italians and Puerto Ricans), some commentators^{9,10} were carried away by hyperbole — asserting that the ceiling principle assumes that “the culprit might be . . . a Lapp for one allele and a Hottentot for the other”. However, if Lapp and Hottentot are replaced by Italians and Puerto Ricans, the assumption is perfectly reasonable. Indeed, it is unreasonable to assume that such genotypes don't occur in the population.

(4) **The ceiling principle is so conservative that it hampers the courtroom application of DNA fingerprinting.** Despite fears that an ultraconservative standard would clip the wings of the DNA fingerprinting, published analyses by several groups^{13,14} agree that the effect is modest: Whereas the product rule typically gives four-locus genotype frequencies of about 10^{-8} - 10^{-9} , the ceiling principle pares them back to about 10^{-6} - 10^{-7} . That extreme assumptions have so little effect only underscores the power of DNA typing and the wisdom of taking a conservative approach.

(5) **The ceiling principle is not actually guaranteed to be conservative.** This concern appears in only a single paper¹⁵, in which the authors point out that the conservativeness of the ceiling principle depends on the component populations $P_1, \dots, P_n$ being themselves well-mixed. The authors take great pains to construct counterexamples in the event that the populations are themselves substructured. In fact, the NRC committee considered this point to be self-evident (although failed to state it clearly enough), as can be seen from the trivial observation that the ceiling principle has no effect on genotype frequencies when

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Opting for conservative procedures — FBI serology laboratory in Washington, DC.

one combines two identical, but substructured populations. In applying the ceiling principle in practice, the committee was confident that any residual substructure in the component subpopulations could safely be ignored in the context of such a conservative scheme.

(6) **The NRC report is causing DNA fingerprinting cases to be thrown out of court⁸.** To the contrary, most courts have used the NRC report as strong evidence that, notwithstanding disagreement over the best solution, there is at least one approach that is indisputably conservative. Of the few cases in which DNA evidence has ever been rejected on population-genetic grounds, virtually all involved evidence predating the NRC report. These courts have cited the report solely for its acknowledgement that a controversy existed and was a reason for constituting the committee.

The NRC report, to be sure, has important flaws. The ceiling principle was not elegant solution, but simply a practical way to sidestep a contentious and unproductive debate. The report had more than its share of miswordings, ambiguities and errors, many of which have been corrected by a vigilant commentator¹⁶. A few poorly worded sentences have been seized upon by lawyers trying to undermine the straightforward calculation of ceiling frequencies (although such arguments have not succeeded). Most important, the report failed to state clearly enough that the ceiling principle was intended as an ultra-conservative calculation, which did not bar experts from providing their own ‘best estimates’ based on the product rule. This failure was responsible for the major misunderstanding of the report. Ironically, it would have been easy to correct.

A law-enforcement perspective

Even as academics debated fine points, forensic scientists got on with their busi-

ness. The FBI and TWGDAM found the ceiling principle to be unnecessarily conservative, but nonetheless promptly adopted precise guidelines for implementing the ceiling principle, correctly clarifying a few ambiguous statements in the NRC report, such as which population databases to include and whether to sum adjacent bins in a frequency distribution. Forensic labs adopted a two-tiered approach, in which experts are prepared to quote both their best estimate and the conservative ceiling bound. As new population-genetic issues arise (such as how to modify or replace the ceiling principle to accommodate the less polymorphic PCR-based systems), the community is preparing to develop further guidelines. Overall, the system meets the spirit and the letter of the NRC report.

Conservative calculations have had no noticeable impact on the use of DNA evidence. In the vast majority of cases, a jury needs to know only that a particular DNA pattern is very rare to weigh it in the context of a case: the distinction between frequencies of 10^{-4} , 10^{-6} and 10^{-8} is irrelevant in the case of suspects identified by other means.

The FBI has also rapidly carried out population surveys, as recommended by the NRC committee. FBI scientists have studied more than 25 distinct subpopulations, as well as 50 separate samples from the US population¹⁷⁻²¹. The effort has yielded a remarkable database for examining allele frequency variation among ethnic groups. Reassuringly, the observed variation is modest for the loci used in forensic analysis and random matches are quite rare, supporting the notion that the FBI's implementation of the product rule is a reasonable best estimate. Nonetheless, the FBI has taken the scientifically sound position that it remains wiser to study new loci empirically than to assume that significant variation can never occur.

Most important, the admissibility of such DNA evidence prepared in accordance with the NRC recommendations is firmly established in virtually all US jurisdictions. In a few, the appellate courts have yet to rule formally, but there is little doubt that they will find such evidence acceptable.

Modest proposals

Although the system ain't broke, there is no shortage of proposals about how to fix it. Some academic commentators advocate a return to the product rule; others propose an approach based on the kinship statistic F_{ST} ; and still others recommend an approach involving likelihood ratios that combine gel electrophoresis artefacts and population-genetic considerations into a single statistic^{9,10,22-24}. Some seek to determine genotype frequencies exactly, while others prefer conservative estimates. The NRC — at the urging of the

National Institute of Justice, representing the academic wing of forensic scientists — has concluded that the best solution is to constitute another *ad hoc* committee on DNA fingerprinting, composed primarily of statisticians and population geneticists.

It is easy to forget that this new debate is purely academic. The most extreme positions range over a mere two orders of magnitude: whether the population frequency of a typical four-locus genotype should be stated, for example, as 10^{-5} or 10^{-7} . The distinction is irrelevant for courtroom use.

Rehashing issues may be a harmless pastime in the academic world, but not so in a legal system that lives by the dictum *stare decisis* (let the decision stand). From the standpoint of law enforcement, it is better to have a settled, if slightly imperfect, rule than ceaselessly to quest after perfection. Already the NRC's intention to re-examine forensic DNA typing has been seized upon by some lawyers as evidence that there remain fundamental problems.

Ad hoc committees typically imagine that they will be able to accomplish their mandate with speed and finality. The original NRC study was anticipated to take one year, but required three. The idea of a second NRC panel was first floated in June 1993 with the optimistic projection that it could report by the end of that year. In fact, the committee has only just begun meeting and will probably not issue a report before late 1995. Even then, any recommendations will take 3 years to ripple through the legal system — guaranteeing that finality will not be achieved on these issues before early 1999. Despite the committee's best efforts, any new report will probably offer new opportunities for misunderstanding that will become apparent only after the panel is disbanded. And, if the new report endorses a different standard, some attorneys are sure to argue, rightly or wrongly, that differences between the reports demonstrate a lack of scientific consensus. These observations are not meant to dissuade the new NRC committee from its mission, but rather to point out the challenge facing any *ad hoc* group.

A sounder approach

The real solution is to recognize that forensic DNA typing has become a mature field and requires a more systematic approach. The NRC report anticipated that rapid evolution of technology would pose a steady succession of questions requiring attention. Its most important recommendation was the establishment of a permanent national committee on forensic DNA typing (NCFDT) to address issues as they arose. If such a committee had been appointed in 1992, it could have made short work of the population-genetics issues, by clarifying, changing or

discarding the original NRC recommendations.

It is encouraging that this NRC recommendation has recently been adopted. The newly enacted DNA Identification Act of 1994 mandates the FBI to establish a DNA advisory board to recommend standards for laboratory procedures, quality assurance and proficiency testing. The act requires open meetings and broad representation, including molecular and population geneticists not affiliated with forensic laboratories, with board nominations to come from professional organizations including the National Academy of Sciences. Ideally, the panel will provide a forum for weighing important issues, including new laboratory techniques, population genetics and proficiency testing.

Most of all, the public needs to understand that the DNA fingerprinting controversy has been resolved. There is no scientific reason to doubt the accuracy of forensic DNA typing results, provided that the testing laboratory and the specific tests are on a par with currently practiced standards in the field. The scientific debates served a salutary purpose: standards were professionalized and research stimulated. But now it is time to move on. □

Eric S. Lander is in the Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, Massachusetts 02142, USA and the Department of Biology, MIT, Cambridge, Massachusetts 02139, USA. Bruce Budowle is in the Forensic Science Research and Training Center, FBI Laboratory, FBI Academy, Quantico, Virginia 22135, USA.

1. Lander, E. S. *Nature* **339**, 501-505 (1989).
2. National Research Council DNA Technology in Forensic Science (National Academy Press, 1992).
3. Lewontin, R. C. & Hartl, D. L. *Science* **254**, 1745-1750 (1991).
4. Chakraborty, R. & Kidd, K. K. *Science* **254**, 1735-1739 (1991).
5. *Science* **266**, 201-203 (1994).
6. Weir, B. S. *Proc. natn. Acad. Sci. U.S.A.* **89**, 11654-11659 (1992).
7. Aldous, P. *Science* **259**, 755-756 (1993).
8. Devlin, B., Risch, N. & Roeder, K. *Science* **259**, 748-749 (1993).
9. Morton, N. E., Collins, A. & Balasz, I. *Proc. natn. Acad. Sci. U.S.A.* **90**, 1892-1896 (1993).
10. Morton, N. E. *Genetica* (in the press).
11. Schellenberger, G. D. et al. *Science* **258**, 668 (1992).
12. Terwilliger, J. D. & Ott, J. *Handbook of Human Genetic Linkage* (Johns Hopkins Univ. Press, 1994).
13. Duncan, G. T., Noppinger, K. & Tracy, M. *Genetica* **88**, 51-57 (1993).
14. Krane, D. E. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 10583-10587 (1992).
15. Simowitz, J. R. & Cohen, J. E. *Am. J. hum. Genet.* **53**, 314-323 (1993).
16. Weir, B. S. *Am. J. hum. Genet.* **52**, 437-440 (1993).
17. Federal Bureau of Investigation VNTR Population Data: A Worldwide Study (1993).
18. Budowle, B., Monson, K. L., Giusti, A. M. & Brown, B. L. *J. Forensic Sci.* **39**, 319-352 (1994).
19. Budowle, B., Monson, K. L., Giusti, A. M. & Brown, B. L. *J. Forensic Sci.* **39**, 988-1008 (1994).
20. Budowle, B. et al. *J. Forensic Sci.* (in the press).
21. Huang, N. E. & Budowle, B. *J. Forensic Sci.* (in the press).
22. Morton, N. E. *Proc. natn. Acad. Sci. U.S.A.* **89**, 2556-2560 (1992).
23. Evelt, I. E., Scranage, J. & Pinchin, R. *Am. J. hum. Genet.* **52**, 498-505 (1993).
24. Balding, D. J. & Nichols, R. A. *Forensic Sci. Int.* **64**, 125-140 (1994).

Measuring magnets on a micron scale

A remarkable account of the micro-measurement of single haematite grains provides data that are too late to provide reassurance for palaeomagnetism, but no doubt help tape manufacturers to better things.

THE use of rock magnetism to reconstruct the Earth's magnetic history depends on the assumption that the magnetization of grains of presumably ferromagnetic material in a rock remains stable for millions or even hundreds of millions of years. In the early days of the 1950s, when the first evidence of past reversals of the Earth's magnetic field began to come to light, there was understandably some anxiety that the succession of magnetic reversals might be artefacts of some environmental fluctuation, perhaps a change of temperature.

Of course, the theoretical people were ready with undemanding reassurance. Find a magnetic particle with dimensions comparable with or smaller than a magnetic domain, and you will find it to be remarkably stable under external influences. One guru was E.C. Stoner from the University of Leeds, but the most comforting reassurance of all came from Louis Néel of Grenoble, who cut the figure of a provincial bourgeois professional who might just as easily have been a banker as a physicist.

The argument was that a sufficiently small particle might be magnetized in one direction or the opposite, that the two states would be separated from each other by an energy barrier and that, if the barrier height were known, the likelihood of switching would be simply described by an expression of Arrhenius type, an exponential function with exponent E/kT , where E is the energy, T the temperature and k is Boltzmann's constant.

But even the comfort that Néel exuded did not prevent some brave spirits from seeking to isolate magnetic particles from the rocks that they had already measured in the hope of measuring their magnetic properties individually. There were obvious difficulties. The magnetic particles might shatter more (or less) easily than the matrix particles in a grinding machine, selecting them (with a permanent magnet) might make them switch direction and, in any case, they were so small that nobody was clear what should be done with them. Attempts to work out something from microwave absorption failed.

Only now, it seems, have those ambitions been fulfilled, at least to judge from an article by M. Lederman and S. Schultz from the University of California at San Diego, and M. Ozaki from the Yokohama City University (*Phys. Rev. Lett.* **73**, 1986–1989; 1994). What the authors have done is to study the magnetization of γ -haematite (Fe_2O_3) particles that span a fivefold

range of size downwards in size from $0.30\text{ }\mu\text{m}$.

How has that been done? With a magnetic force microscope, or MFM, of course. And what is that? It is an atomic force microscope in which the force is not mechanical, but is generated by a microsolenoid. The haematite particles were selected for measurement from those deposited on the grid of a transmission electron microscope. Both were picked to have large aspect ratios (roughly 5:1) on the grounds that they would then be unambiguous single domains, that the magnetization would be along the major axis of the ellipsoid and that the coercive force would be large.

The coercivity of haematite is indeed very large, under no circumstances less than 900 Oe. Predictions by Stoner in the late 1940s that it is more difficult to reverse the direction of a single magnetized domain by applying a field in the opposite direction are indeed confirmed. Over 90° , the coercive force ranges from roughly 900 Oe to more than 1,500 Oe. The underlying model is that when the external field is applied at a largish angle relative to the ellipsoid axis, the magnetization will rotate, but otherwise it will have to turn back on itself.

But Lederman and his colleagues are seeking to fry other fish. By their account, there has recently been evidence that collections of many small particles do not properly follow the expected Arrhenius law. In particular, for a collection of magnetized particles in an external switching field, the proportion retaining their original direction should be an exponentially decreasing function of t/τ , where t is the time and τ is a characteristic time proportional to $\exp(E/kT)$, and which of course may differ from one particle to another.

Specifically, previous experiments (with several particles) are said to have found that the magnetization of a sample is proportional to the logarithm of the time. The obvious explanation, that the characteristic times τ of the particles are widely distributed over some range, is hardly falsifiable for a collection of particles, given that the distributions of τ may be virtually chosen at will. But the data have also been used to suggest that novel phenomena, such as quantum tunnelling, may be responsible.

So Lederman and his colleagues set out to study the probability of the reversal of the magnetization of single haematite domains by applying a powerful field in one

direction followed by a carefully measured field in the opposite direction for an interval ranging from a few milliseconds to 10 seconds. It is obviously a tedious business; the tip of the MFM has to be withdrawn on each occasion, and the whole observation repeated dozens of times to provide statistics for an estimate of the frequency of switching, or for a single data-point.

How well does it work out? The conclusion is that the Arrhenius law does not apply. (Presumably an article demonstrating that it did would "not have succeeded in the competition for space", as the euphemism has it.) But the interpretation offered is of more than passing interest.

Comparing the variation of switching probability at different switching fields and different times, the authors conclude that a field in direct opposition to the magnetization of a single domain will nucleate magnetization in its own direction at several points. Then, the argument goes, the likelihood that the whole domain will switch will be determined not by a single barrier energy, but by the chance that geometrically separate nuclei will afterwards grow together. Newly named phenomena are not required.

Many will recognize in that statement a neat problem waiting for a solution, that of calculating the probability that two or even several nuclei of nucleation points in a domain will come to determine the magnetization of a whole magnetic domain. But it may be too late to summon a graduate student to carry out the task; by all accounts, numerical solutions are already under way.

For what it is worth, nowhere in the article do Lederman and his colleagues refer to the problems of the pioneer palaeomagnetism community, but that is hardly surprising. For the truth is that the old worry has been exorcised by the widespread recognition that reversals of the Earth's magnetic field do actually occur, and that the magnetized particles in rocks keep a largely faithful record of them.

And, of course, the people with a now-vivid interest in the probability that an external field will switch the direction of magnetization are the manufacturers of magnetic recording tape, forever seeking fidelity in recording. The great Néel, for all his style and foresight, can hardly have foreseen how the demands of technology would belatedly drive experimenters into unearthing data that would have been invaluable 40 years ago.

John Maddox

The little Big Bang

Gregory A. Petsko

It is the biophysical equivalent of the Big Bang: a flash of light ruptures the covalent bond between carbon monoxide and iron in the protein myoglobin. The events following this submicroscopic explosion have been the subject of as much speculation as — and a lot more experimentation than — its cosmological counterpart. Now, in papers on page 808 of this issue¹ and the October number of *Nature Structural Biology*², two groups present atomic-resolution structures of the photolysed state. They have been obtained by a new technique, ultra-low-temperature protein crystallography.

Myoglobin is the hydrogen atom of biology. It was the first protein to have its three-dimensional structure determined, and since then biophysical chemists and structural biologists have done everything to it but split its atoms. Myoglobin is a small, mostly helical, monomeric oxygen storage protein. Because it serves as a reserve supply of oxygen in vertebrate skeletal muscle, myoglobin is particularly abundant in deep-diving mammals such as whales and seals. Its three-dimensional structure is very similar to that of both chains of the blood-oxygen transport protein haemoglobin, which has sometimes

been called the hydrogen molecule of biology. Oxygen (or, here, carbon monoxide) binds to these proteins by formation of a coordinate covalent bond with an iron atom located near the centre of the protein. The iron is part of an organic prosthetic group, the haem, the geometry of which changes with ligand binding. This change is communicated to the protein, which changes its conformation in response through a histidine residue bonded to the iron (the 'proximal' histidine) and through nonbonded contacts.

What makes myoglobin the perfect model system for biophysics is that these changes can be triggered by light. The energy from individual photons that gets deposited in the haem initiates a 'proteinquake'³, breaking the bond between the haem iron and the ligand, and initiating a sequence of conformational rearrangements in the protein. Once dissociated, the carbon monoxide has three fates: it can rebind immediately, it can rattle around the haem pocket for a while and then rebind, or it can escape through the matrix of protein. Because this last possibility requires that protein atoms move in order to open a transient pathway for the ligand to escape, studies of the photolysed state of the myoglobin-carbon monoxide (Mb-CO) complex helped spawn the

field of protein dynamics⁴. At liquid-helium temperatures, there is not enough thermal energy for CO to leave the protein or to rebind rapidly, and the unbound state can be trapped long enough for kinetic and structural observation (part *a* in the figure).

Yet while forests have gone to the blade supplying paper for studies, spectroscopic and otherwise, on the 'simple' reaction

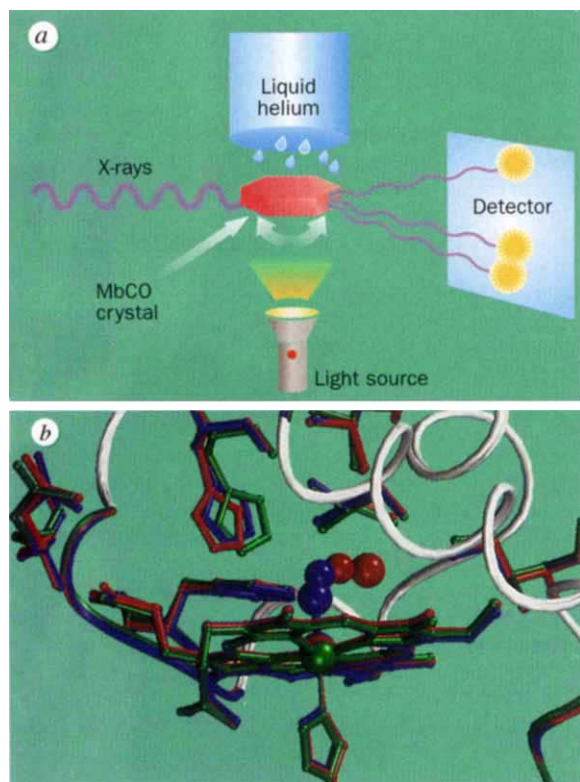


many fundamental questions remain. Not least are these: where does the CO go, and how does the protein respond? These two questions have now been answered by crystal structures^{1,2} of Mb*CO, the intermediate state with the Fe-CO bond broken but with the ligand still inside the protein.

Although Schlichting *et al.*¹ and Teng *et al.*² arrive at somewhat differing results, both studies represent astonishing technical achievements. They achieve a new low in protein crystallography: structure determination at liquid-helium temperatures⁵. Thermally activated kinetic processes that are too fast to catch even at 80 K (ref. 6) are now within the grasp of the kinetic crystallographer. At liquid-helium temperatures, states separated by activation enthalpy barriers as low as 5 to 0.5 kcal mol⁻¹ are stable for long enough to allow rapid collection of X-ray diffraction data. Such intermediates are of great importance in bacteriorhodopsin and the photosynthetic reaction centre, as well as in many enzymatic reactions.

Although cryocrystallography and monochromatic radiation from a synchrotron source are at the heart of the technical approaches used by both groups, there are several differences. Schlichting *et al.* conducted their experiments on a mutant myoglobin that crystallizes in a hexagonal lattice, employed continuous photolysis with a white-light source at very low power, and collected data at a temperature of 20 K. Teng *et al.* used the monoclinic form of native sperm-whale myoglobin, short photolysis by a more intense laser source, and a temperature of 40 K. At this higher temperature, 15–20 per cent of the photolysed CO ligands will rebind over the time required to collect a diffraction pattern. Spectroscopic characterization of their crystal also showed a significant amount (30 per cent) of the oxidized form of the protein, which has a water molecule bound tightly to the haem iron.

Both groups find that the CO has moved away from the iron and that the iron has moved out of the mean haem plane. Both see a definite orientation for the C-O axis in the pocket, but neither can discriminate between the two possibilities for which end is 'C'. Schlichting *et al.* find the CO lying flat along the haem plane at



a, Experimental protocol for obtaining the crystal structure of photolysed MbCO. An optically thin MbCO crystal is oscillated in the crossfire of a beam of monochromatic X-rays from a synchrotron that stimulates diffraction, a beam of visible light from a laser or conventional light source that zaps the CO off the haem iron, and a stream of helium gas/liquid that keeps the crystal at a temperature below 40 K to slow down the rebinding reaction. A few minutes' exposure is all that is required to collect the diffraction pattern. *b*, Changes in the active-site geometry among the CO-bound (blue), photolysed (red) and unligated (green) states. Both Schlichting *et al.*¹ and Teng *et al.*² see significant changes upon photolysis in the position of the CO and the out-of-haem plane distance of the iron (large spheres show 30 per cent of van der Waals' radii for these atoms). Schlichting *et al.* report changes upon photolysis in the position of the distal His (upper left), Arg CD3 (extreme left), proximal His (lower centre) and F helix, as illustrated here.

van der Waals' distance from it, with the nearer end 3.6 Å away from the iron, which has moved 0.2 Å out of the haem plane. Teng *et al.* show the CO tilted in a way more like the CO-bound form, with the nearer end much closer to the iron which has moved by less than 0.1 Å. They find no significant changes in the protein, whereas Schlichting *et al.* describe motions of several groups near the distal haem pocket and a partial relaxation of the proximal histidine (part *b* in the figure). It is unlikely that the discrepancy is due to crystal-packing forces — there are too many spectroscopic and kinetic similarities between the two crystal forms and solution samples. Moreover, the conformations of the protein backbones are virtually identical in the two crystal systems⁷, and the mutation used to produce the hexagonal crystal form is on the surface and doesn't affect the properties of the protein. At least some of the differences between the two sets of results probably arise from the presence of oxidized Mb and rebound CO in the experiment of Teng *et al.*; variations in photolysis protocol may account for the rest.

Schlichting *et al.* find the CO in the photodissociated state at a site where a water molecule sits in the unligated state. Because there isn't room for both the water and the ligand in the pocket, the question arises as to how the water gets out so that a ligand can get in. Perhaps fluctuations in the protein structure pop the distal histidine out of the pocket, dragging the water along with it. The reverse process could allow oxygen in from the solvent. The outwards motion of the distal histidine upon photolysis, which is presumably due to electrostatic interaction between the ligand and the histidine N-epsilon, is something of a surprise. This interaction, unfavourable for CO, would presumably be favourable for oxygen and may account for the difference in binding energy seen between CO and oxygen intermediates⁴. The changes seen upon

photolysis in the iron out-of-plane distance and in the proximal histidine by Schlichting *et al.* are consistent with what was expected from spectroscopy^{8–10} and molecular dynamics calculations^{11,12}. In fact, the agreement is quantitative for the iron movement, a remarkable testimony to the power of all these methods. Even simple computations are able to predict the ligand position surprisingly well¹³ — most reassuring, because this work is one of the few experimental tests of prediction from simulation.

The atomic coordinates from these studies provide a structural basis for the interpretation of many past and future experiments on haem protein biophysical

chemistry. They also serve as the starting point for a plethora of new computational studies. Aficionados of the photosynthetic reaction centre and other light-activated proteins will no doubt be among the first to make use of the possibilities opened up by this new technology. Meanwhile, biophysicists can enjoy something their cosmologist colleagues can only dream of: actual pictures of their counterpart to Steven Weinberg's fabled First Three Minutes. □

Gregory A. Petsko is at the Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, Massachusetts 02254, USA.

HUBBLE CONSTANT

The Universe in crisis

George H. Jacoby

Judge not the play before the play is done.
(Francis Quarles)

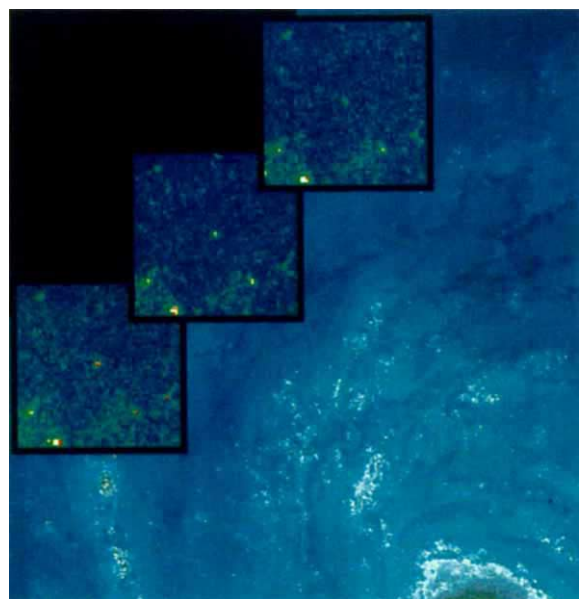
TECHNOLOGY has caught up with Edwin Hubble's question of the 1920s: what are the scale and age of the Universe? Preliminary answers given last month by Pierce *et al.*¹, and strongly backed up on page 757 of this issue by Freedman *et al.*², indicate that the Universe is younger than its oldest stars, an apparent impossibility that will force a re-examination of our Universe model and how stellar ages are measured.

At the end of last month Pierce *et al.*¹ announced the first direct measurement of the distance to the galaxy NGC4571 in the Virgo cluster using Cepheid variable stars (see the News and Views by Hogan³). This distance implies that the Hubble constant H_0 , the ratio of the recession velocity to the distance for a galaxy, is $87 \pm 7 \text{ km s}^{-1} \text{ Mpc}^{-1}$. In the absence of accelerating forces, the age of the Universe is less than $1/H_0$ and, in the standard model of the Big Bang, is $2/3 \times 1/H_0$, or 7×10^9 years. In contrast, some stars are thought to be 16×10^9 years old. If H_0 is correct, then alleviating this age 'crisis' demands that at least one of the following be considered: stellar ages are too high, an accelerating force exists, or the standard, closed-Universe, Big Bang model is incorrect.

The effort by Pierce and collaborators is the first scene in the final act of the decades-long drama of the extragalactic distance scale.

Over the years, key players have made claims and counter-claims that, as in any good mystery story, were later found to be riddled with false clues.

However, the story doesn't end with NGC4571, as it is only one galaxy of hundreds in the Virgo cluster. All those galaxies may not be at the same distance. Moreover, the observations made by Pierce and colleagues suffer from the limitations of being made from the ground. Despite the excellent conditions and advanced technology cameras on Mauna Kea in Hawaii, looking through the Earth's atmosphere blurs stellar images. Faint stars may be blended into the Cepheids to make them appear brighter and closer than they would under finer



Visual-band image of the outer spiral structure in the Virgo galaxy M100, and (inset) successive images of one of the Cepheid variable stars, taken by the HST.

W.L. Freedman & L. Ferrarese/NASA

- Schlichting, I., Berendzen, J., Phillips, G. N. Jr & Sweet, R. M. *Nature* **371**, 808–812 (1994).
- Teng, T.-Y., Šrajer, V. & Moffat, K. *Nature Struct. Biol.* **1**, 701–705 (1994).
- Ansari, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 5000–5004 (1985).
- Austin, R. H., Beeson, K. W., Eisenstein, L., Frauenfelder, H. & Gunsalus, I. C. *Biochemistry* **14**, 5355–5373 (1975).
- Teng, T.-Y., Schildkamp, W., Dolmer, P. & Moffat, K. *J. appl. Crystallogr.* **27**, 133–139 (1994).
- Hartmann, H. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 4967–4971 (1982).
- Phillips, G. N. Jr *Biophys. J.* **57**, 381–383 (1990).
- Della Longa, S., Ascone, I., Fontaine, A., Congiu Castellano, A. & Bianconi, A. *Eur. Biophys. J.* (in the press).
- Dasgupta, S. & Spiro, T. G. *Biochemistry* **25**, 5941–5948 (1986).
- Chance, M. R., Parkhurst, J. L., Woolery, G. L. & Chance, B. *J. biol. Chem.* **261**, 5689–5692 (1986).
- Gibson, Q. H., Regan, R., Elber, R., Olson, J. S. & Carver, T. E. *J. biol. Chem.* **268**, 17908–17916 (1992).
- Kuczera, K., Lambry, J. C., Martin, J.-L. & Karplus, M. *Proc. natn. Acad. Sci. U.S.A.* **90**, 5805–5807 (1993).
- Sassaroli, M. & Rousseau, D. L. *J. biol. Chem.* **261**, 16292–16294 (1986).

inspection. Weather and scheduling constraints mean that the light curves of the Cepheids cannot be sampled ideally, thereby complicating the determination of periods and distances. Furthermore, Pierce *et al.* found only three Cepheids. Critics of the high Hubble constant demand more proof. Given the history of this field and its relevance to Big Bang theory³, more proof seems prudent.

Now, scene two of the final act has unfolded. Freedman *et al.*² report that the Hubble Space Telescope (HST) has achieved a fundamental design goal, one that Hubble himself would surely have loved to have witnessed: the undeniable identification of 20 Cepheid variables in M100, a beautiful spiral galaxy in Virgo. Critics cannot now argue that there are too few Cepheids, nor, with the superb images of the HST, can they protest about blended images. With the guaranteed 'clear' weather in space, the Cepheid cyclical variations are perfectly sampled; the periods are correct. Thus, the distance to M100, 17.1 ± 1.8 Mpc, only 15 per cent farther than NGC4571, cannot be questioned. Given this distance, Freedman *et al.* find $H_0 = 80 \pm 17$ km s⁻¹ Mpc⁻¹, implying an age for the Universe of 8×10^9 years in the standard Big Bang model. The conflict with the stellar ages persists.

The tremendous significance of their announcement lies in the use of Cepheids to derive an unquestionable distance to a Virgo galaxy, and in demonstrating that HST now makes easy what was impossible a few years ago. It also bolsters the high H_0 found by Pierce *et al.* and validates the results from less direct methods.

We now know the distance to two galaxies in Virgo. If both happen to be in front of the cluster, then the bulk of Virgo galaxies might be more distant and H_0 could still be small. The evidence, however, is that these galaxies are near the centre of Virgo. In fact, they have long been considered representatives of the Virgo cluster⁴.

Nevertheless, future work will address the depth issue; for now, it dominates the uncertainty in H_0 from Virgo Cepheids. To be certain, the HST team plans to obtain distances to two more galaxies in Virgo, two in Fornax (a southern galaxy cluster), and about ten more galaxies to serve as check points for other techniques (surface brightness fluctuations, planetary nebulae, globular clusters, type II supernovae and the Tully-Fisher relation).

All of the above techniques have been used previously to derive the Virgo distance, but those methods are not perceived to be as reliable as Cepheids because they require an intermediate calibration step, usually based on Cepheids in nearby galaxies. These methods have been refined over the years to a good level of accuracy and are now

vindicated by the new results; H_0 from most falls between 70 and 85 km s⁻¹ Mpc⁻¹ (ref. 5).

One method, though, stands apart: distances derived from type Ia supernovae yield $H_0 = 54 \pm 8$ (ref. 6). An HST team led by Allan Sandage is pursuing additional distant check points to verify the accuracy of supernovae distances. Others^{7,8} claim that type Ia supernovae require more parameters in their calibration, which when applied yield $H_0 = 67 \pm 7$ km s⁻¹ Mpc⁻¹. The focus of the distance-scale debate has recently centred here — what is the proper calibration for type Ia supernovae? With the Virgo Cepheid results, the question is in the spotlight again.

The next act in the play will concentrate on resolving this final conflict between results. The final scenes will address the possibility that systematic errors exist. For example, how does the different chemical composition of the Cepheids in different galaxies affect our estimates of their distances? The preponderance of clues at this time favours a large H_0 and a young

Universe. If the clues are being interpreted correctly, the curtain soon will fall on the H_0 controversy, and the audience will file out in anticipation of the sequel. That will focus on how theorists resolve the age crisis.

We live in a special time: after millennia of not knowing the size and age of our Universe, we soon will. We also live in a time of crisis, for we may be forced to accept something new about the ages of the stars or the nature of the Universe. □

George H. Jacoby is at the National Optical Astronomy Observatories, 950 North Cherry Avenue, PO Box 26732, Tucson, Arizona 85726, USA.

1. Pierce, M. J. *et al.* *Nature* **371**, 385–389 (1994).
2. Freedman, W. L. *et al.* *Nature* **371**, 757–762 (1994).
3. Hogan, C. J. *Nature* **371**, 374–375 (1994).
4. Sandage, A., Binggeli, B. & Tammann, G. A. *Astr. J.* **90**, 395–404 (1985).
5. Jacoby, G. H. *et al.* *Publ. astr. Soc. Pac.* **104**, 599–662 (1992).
6. Saha, A. *et al.* *Astrophys. J.* (in the press).
7. Hamuy, M. *et al.* *Astr. J.* (in the press).
8. Riess, A., Kirshner, R. & Press, W. *Astrophys. J.* (in the press).

CELL CYCLE

Ubiquitin with everything

Jonathon Pines

Is the cell cycle simply a series of degrading events? Certainly, the destruction of mitotic cyclins is already strongly implicated in controlling the exit from mitosis into the next interphase. Now a paper by Schwob and colleagues, published in last week's *Cell*¹, suggests that the destruction of a cyclin-dependent kinase inhibitor is essential for the initiation of DNA replication. Thus it looks as if the phases of the cell cycle are coordinated through the programmed destruction of cell-cycle regulators. Furthermore, ubiquitin-dependent proteolysis is responsible for the destruction of the regulators of both DNA replication and of mitosis.

The cell cycle is thought to be controlled by the sequential activation of a series of cyclin-dependent kinase (CDK) complexes, which regulate the various checkpoints controlling DNA replication and mitosis. In yeast, one CDK (Cdc28 in budding yeast, Cdc2 in fission yeast) forms a succession of complexes with different cyclins during the cell cycle. In budding yeast the major G1 phase checkpoint of the cell cycle, Start, is regulated by the G1 cyclins (Clns). Cln1 and Cln2 disappear as cells replicate their DNA. The latter part of the cell cycle is controlled by the B-type cyclins (Clbs): the S-phase (Clb5 and Clb6) and mitotic (Clb1–4) cyclins. Thus one of the most fundamental questions in cell-cycle regulation concerns the nature of the mechanisms that ensure the correct alternation of the different cyclin-CDK

complexes. Recent work, primarily from Nasmyth and colleagues, suggests that in large part the answer to this problem lies in cell-cycle-dependent alternating states of stability and instability of the cyclins².

Cln1 and Cln2 are constitutively unstable throughout the cell cycle, so their levels are determined by the rate of their transcription. Cln1 and Cln2 transcription is initiated in G1 phase once cells have reached a critical size, probably by the Cln3-Cdc28 kinase complex (for review see ref. 3). Cln1 and Cln2 are required for the cell to pass Start and begin DNA replication, and their transcription is turned off when Clb1- and Clb2-Cdc28 protein kinase activity appears in G2 phase⁴. Clb1 and Clb2 only appear in G2 phase, in part because they are only transcribed once cells have passed Start and begun DNA replication. The other important factor is that they are unstable throughout G1 phase, and the Cln1- and Cln2-Cdc28 kinases are required to turn off the proteolysis machinery².

So far, so logical — Clb1 and Clb2 can only accumulate after Cln1 and Cln2 become active, and Clb1 and Clb2 in turn down-regulate Cln1 and Cln2. The problem was that Clb5 and Clb6, which are needed for efficient DNA replication, accumulate with the same kinetics as Cln1 and Cln2 (refs 5, 6). What then prevents the premature activation of this first set of Clbs?

Schwob *et al.*¹ show that the answer is a NATURE • VOL 371 • 27 OCTOBER 1994

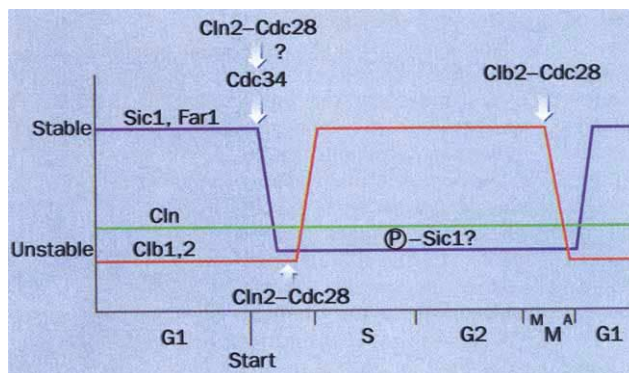
specific Clb–Cdc28 inhibitor protein, p40^{Sic1} (Sic1). Sic1 accumulates in G1 phase, and binds and inhibits the Clb5–Cdc28 complex; it does not affect the activity of the Cln–Cdc28 complexes. Once cells have passed Start, Sic1 is degraded by a ubiquitin-dependent proteolysis pathway involving the Cdc34 ubiquitin-conjugating enzyme. The Sic1-destruction machinery remains active throughout S and G2 phases, and is finally inactivated at mitosis. Yeast cells mutant in Cdc34 arrest in G1 phase and are unable to begin DNA replication. The deletion of Sic1 partially rescues a Cdc34 mutant cell; the cell is able to replicate its DNA, albeit more slowly than normal, but arrests before mitosis. So although there must be other Cdc34 substrates later in the cell cycle, the main S phase inhibitor is Sic1.

There may be a family of cell-cycle regulators degraded in the same way as Sic1. The G1 cyclin–CDK inhibitor, Far1 (ref. 7), has very similar stability to Sic1, and its degradation is dependent on Cdc34 (ref. 8). The destruction pattern of Sic1 and Far1 is the mirror image of that of Clb1 and Clb2, the degradation of which is triggered at mitosis and inactivated once cells pass Start. Clb1 and Clb2 are also degraded by a ubiquitin-dependent pathway, but through a different ubiquitin-conjugating enzyme. Clb1 and Clb2 could be homologues of the metazoan B-type cyclins, which trigger their own destruction when they activate Cdc2 at mitosis⁹. Thus Clb1 and Clb2 might activate their own destruction, and simultaneously stop destruction of Sic1 and Far1 (see figure).

Several issues are raised by all these findings. First, is this an evolutionarily conserved mechanism? The fission yeast Rum1 CDK inhibitor¹⁰ is also able to delay the onset of S phase, but Sic1 and Rum1 may not be homologues because high levels of Rum1 cause cells to re-replicate their DNA¹⁰. Mammalian cells have a Cdc34 homologue¹¹, so there might be a mammalian Sic1 involved in the regulation of S phase. Several mammalian CDK inhibitors have been isolated, but none of them has extensive structural homology to Sic1, nor do any of them display a similar behaviour pattern through the cell cycle. One inhibitor, p16 (ref. 12), binds exclusively to the D-type cyclin partner CDK4, and thus resembles Far1 in binding only to the G1 cyclin–CDK complexes. Two other mammalian inhibitors, p21 and p27, seem to be involved in the cellular response to extrinsic factors, such as DNA damage¹³ or nega-

tive growth factors¹⁴, rather than coordinating the cell cycle *per se*¹⁵. Thus the isolation of mammalian homologues of Sic1 and Rum1 is eagerly anticipated.

Most intriguing of all is how ubiquitin-dependent proteolysis is able to degrade one set of proteins only in mitosis and G1 phase, and a completely different set of proteins only in S and G2 phases. It is likely that there is regulation at the level of both the substrate and the ubiquitin-



Representation of the alternating periods of stability of the cyclins and CDK inhibitors in the budding yeast cell cycle (note that this does not show the relative levels of the proteins). The effect of transcription means that Cln1 and Cln2, which are rapidly turned over throughout the cell cycle, peak at Start, and that Clb1 and Clb2 do not begin to accumulate until G2 phase. The possible influences of Cln2–Cdc28 phosphorylation on Cdc34 activation, on marking Sic1 as a proteolytic substrate, and in turning off Clb proteolysis, are also shown.

conjugating machinery. Schwob *et al.* observed that Sic1 may be a substrate of Cln2–Cdc28, and this phosphorylation might target the protein for destruction. Perhaps only the phosphorylated form of Sic1 is recognized by the ubiquitin-conjugating enzyme or perhaps this form of Sic1 is transported to the proteolytic machinery.

Several different ubiquitin-conjugating enzymes are known (at least nine in yeast), and these may recognize particular motifs shared by a set of proteins. One candidate motif is the 'destruction box', which is shared by all mitotic cyclins isolated to date, including Clb1 and Clb2. The destruction box is necessary for cyclin ubiquitination, and almost certainly marks cyclins for degradation at mitosis. The destruction box might be recognized

by a particular ubiquitin-conjugating enzyme or enzymes, and an apparently cyclin-specific one has been isolated from clam oocytes¹⁶. Just how specific any such enzyme will be for a particular motif is unknown. For example, there could be a cyclin A-specific and a cyclin B-specific enzyme because the sequence of the destruction box varies between the A- and B-type cyclins, and the A-type cyclins are destroyed in metaphase whereas the B-types are destroyed when cells enter anaphase. It will be interesting to see whether a common structural motif is found in Far1 and Sic1 and other proteins that are degraded in a Cdc34-dependent manner as cells pass Start — or, indeed, between other proteins that are degraded in mitosis, such as CENP-E (ref. 17) and the components that link sister chromatids together¹⁸.

Finally, the ubiquitin-conjugating machinery itself may be activated only at particular points in the cell cycle, and naturally this could best be achieved through phosphorylation by cyclin–CDK kinases. Indeed, Cdc34 is a phosphoprotein¹⁹, and in a re-constituted proteolysis system from clam oocyte extracts, a novel ubiquitin ligase enzyme

required for cyclin B ubiquitination is only active in mitotic extracts, but can be stimulated in interphase by Cdc2 kinase¹⁶. This could be the level at which the B-type cyclins trigger their own destruction, and (by invoking allostery) at which the G1 cyclin kinases turn it off again.

The equal and opposite effects of the G1 cyclin kinases and the G2 cyclin kinases are also exerted on cell-cycle transcription factors in yeast⁴. It will be fascinating to discover the molecular mechanisms by which two different kinase complexes, using the same catalytic subunit, may potentially have diametrically opposed effects on the same substrate. □

Jonathon Pines is at the Wellcome/CRC Institute, Tennis Court Road, Cambridge CB2 1QR, UK.

- Schwob, E., Böhm, T., Mendenhall, M. D. & Nasmyth, K. *Cell* **79**, 233–244 (1994).
- Amon, A., Imrigger, S. & Nasmyth, K. *Cell* **77**, 1037–1050 (1994).
- Koch, C. & Nasmyth, K. *Curr. Opin. Cell Biol.* **6**, 451–459 (1994).
- Amon, A., Tyers, M., Futcher, B. & Nasmyth, K. *Cell* **74**, 993–1007 (1993).
- Epstein, C. B. & Cross, F. *Genes Dev.* **6**, 1695–1702 (1992).
- Schwob, E. & Nasmyth, K. *Genes Dev.* **7**, 1160–1175 (1993).
- Chang, F. & Herskowitz, I. *Cell* **63**, 999–1011 (1990).
- McKinney, J. D., Chang, F., Heintz, N. & Cross, F. R. *Genes Dev.* **7**, 833–843 (1993).
- Luca, F. C., Shibuya, E. K., Dohrmann, C. E. & Ruderman, J. V. *EMBO J.* **10**, 4311–4320 (1991).

- Moreno, S. & Nurse, P. *Nature* **367**, 236–242 (1994).
- Plon, S. E., Leppig, K. A., Do, H.-N. & Groudine, M. *Proc. natn. Acad. Sci. U.S.A.* **90**, 10484–10488 (1993).
- Serrano, M., Hannon, G. J. & Beach, D. *Nature* **366**, 704–707 (1993).
- Dulic, V. *et al.* *Cell* **76**, 1013–1023 (1994).
- Polyak, K. *et al.* *Genes Dev.* **8**, 9–22 (1994).
- Pines, J. *Trends biochem. Sci.* **19**, 143–145 (1994).
- Hershko, A. *et al.* *J. biol. Chem.* **269**, 4940–4946 (1994).
- Brown, K., Coulson, R., Yen, T. & Cleveland, D. J. *Cell Biol.* **125**, 1303–1312 (1994).
- Holloway, S. L., Glotzer, M., King, R. W. & Murray, A. W. *Cell* **73**, 1393–1402 (1993).
- Goehl, M., Goetsch, L. & Byers, B. *Molec. cell. Biol.* **14**, 3022–3029 (1994).

A Heisenberg microscope

Christopher Foot

ATOM manipulation using laser beams is starting to allow experiments to be performed which until recently one would have considered only as 'gedanken' or thought experiments. A prime example is the 'Heisenberg microscope' which Heisenberg used to illustrate his famous Uncertainty Principle back in the earliest days of quantum mechanics. Now, J. Mlynek's group at the University of Konstanz (T. Pfau, S. Spälter, C. Kurtseifer, C. R. Ekstrom & J. Mlynek *Phys. Rev. Lett.* **73**, 1223–1226; 1994) have carried out a beautiful experiment in which they managed to observe the effect of a single spontaneously emitted photon on an atom. By a straightforward extension of this apparatus, they should be able to make measurements similar to Heisenberg's 'gedanken' experiment.

In the original experiment Heisenberg asked the reader to imagine observing a single particle through an idealized microscope (Fig. 1) by observing the light scattered from it, and to consider how accurately its position and momentum could be measured. Optics tells us that, in theory, this apparatus enables the position of a particle to be determined within a region of $\Delta x = \lambda/\sin \theta$, where λ is the wavelength of the scattered light and θ is the half angle of the cone over which the lens collects light ($\sin \theta$ is the numerical aperture).

Next, the momentum. What happens when a scattered photon is observed? Heisenberg was not concerned with the value of the momentum itself but rather the smallest uncertainty with which it

could in principle be measured. Even though the momentum of the photon and particle can be assumed to be accurately known beforehand and momentum is conserved during the scattering, uncertainty will arise because the photon goes in a random direction and the particle recoils in the opposite direction.

The range of angles for which photons are collected is determined by the finite aperture of the lens, so there is an uncertainty in the x component of its momentum of $\Delta p_x = (h/\lambda) \sin \theta$, where h/λ (Planck's constant over the wavelength) is the magnitude of the momentum of the photon. The uncertainty in the particle's final momentum will be the same as for the photon and so the product of the uncertainties for the position and momentum of the particle has a minimum value of roughly Planck's constant: $\Delta x \Delta p_x \geq h$.

This is the Uncertainty Principle. It can be derived directly from basic postulates of quantum mechanics but Heisenberg's example gives an intuitive physical picture of why we cannot know simultaneously precise values of the position and momentum.

The Konstanz experiment is pretty much this textbook description come to life. A well collimated atomic beam is used so that the transverse component of the momentum is well defined. In terms of the geometry of Fig. 1, the atoms have a large momentum in the y -direction, but it is still the component along x that is of interest (Fig. 2). The light in this modern-day experiment is provided by a laser beam through which the atoms pass. Where it differs from the textbooks, though, is that the scattered photons are not detected.

Instead, Mlynek's group was aiming in the present experiment to measure the effect of a single spontaneously emitted photon on the spatial coherence of an atomic beam. This is achieved by sending a beam of metastable helium atoms through a standing-wave light field which acts as a diffraction grating by producing a periodic phase modulation on the transverse atomic wavefunction. In other words, the atoms are treated as matter-waves which are diffracted by a grating made from light — the opposite of the usual situation. This idea of making optics

for atoms is fairly new, and much of the pioneering work in this field has been carried out at Konstanz.

To observe the diffraction patterns a movable detector with a 5-micrometre entrance slit is scanned in the x -direction (see Fig. 2). The separation of the diffraction peaks is inversely proportional to the spacing of the grating formed by the standing wave, and can be varied by changing the angle of incidence, α , of the laser beam on the mirror.

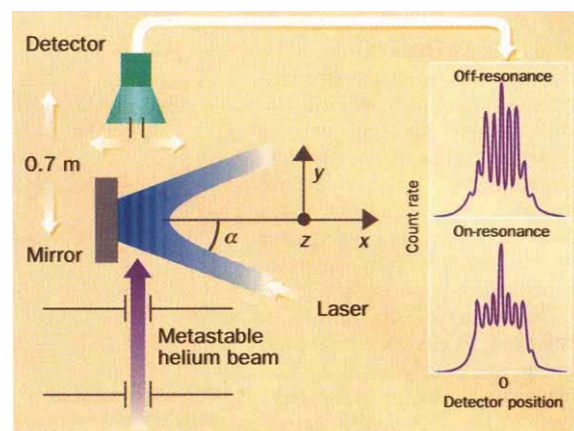


FIG. 2 In the experiment of Pfau *et al.* an atomic beam is passed through a standing wave formed by the reflection of a laser beam at a mirror. The resulting diffraction pattern is recorded by scanning a detector in the x -direction, both when the laser frequency is off- and on-resonance with the atomic transition. On-resonance, the atoms are excited and the random recoil from the spontaneously emitted photons broadens the peaks. The central peak arises from atoms in states not affected by the laser.

Pfau *et al.* record two types of diffraction pattern: first, with the laser slightly off-resonance from an atomic transition; second, with the laser exactly on-resonance. In the first case, the atoms remain in the ground state during diffraction, and the peaks are well defined. In the second, half the atoms are in the excited state, and the recoil of the atoms from the random spontaneous photon emitted when the excited state decays spreads these peaks.

Conditions are carefully controlled so that there is no great probability of more than one spontaneous emission per atom. Nevertheless there is an easily observable effect, especially when peaks are closer together (for large α). The two types of pattern are combined mathematically to extract the pattern for the excited atoms alone.

The usual Fourier transform method of classical optics is then applied to determine the loss of transverse coherence of the atomic beam. However, the quantum nature of a single spontaneous emission means that the atomic momentum can only change by a discrete amount, so the loss of coherence should vary in an oscillatory way — and this is indeed observed. The experimental verification of this

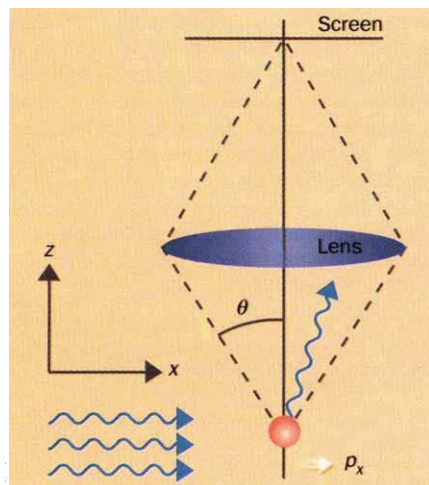


FIG. 1 Heisenberg's microscope was originally a 'gedanken' experiment to illustrate the Uncertainty Principle. When a particle is observed by means of scattered light (or any other method) there is a fundamental limit to how well the position and momentum can simultaneously be known.

effect, which is quite distinct from when there is no quantization, is the significant achievement in this paper.

Extending the apparatus to produce a Heisenberg microscope should be simple: add a lens and a detection system to observe the atoms while they are in the standing wave. By recording the diffraction pattern only for those atoms that emit a photon into the detector's solid angle, the excited pattern could be observed directly. For each of these atoms the spread of momentum, p_x , produced by the recoil from the photon can only be between $\pm (h/\lambda) \sin \theta$, so the peaks would be narrower.

This corresponds very closely to Heisenberg's original idea; by making the aperture ($\sin \theta$) of the lens smaller, the uncertainty in the position of the atom would increase and correspondingly the uncertainty in the momentum (peak width) would decrease. Here, we have a very carefully controlled quantum measurement at the fundamental limit of precision. The founders of quantum mechanics could only have dreamt about this. □

Christopher Foot is in the Clarendon Laboratory, University of Oxford, Parks Road, Oxford OX1 3PU, UK.

VISION

Illusory figures and real neurons

Gerald Westheimer

THERE was a time when psychologists described and measured perceptual phenomena and developed concepts about them, while physiologists recorded action potentials in neurons and proposed schemes of neural circuits. On the rare occasions when the two met, they eyed each other with suspicion.

In the 1860s and 1870s Mach and Hering argued convincingly from psychophysical evidence for the existence of centre/surround opponency in spatial vision. But it took almost 100 years, and a great deal of technical advance, before neurophysiologists could demonstrate that retinal neurons in fact exhibit this property. As a result we can now account nicely for a whole body of visual phenomena, such as Mach bands and contrast induction, in terms of retinal processing. But the retina is an independent neural outpost, connected to the brain by a one-way path. Once the signals reach the brain proper, strong interconnectivity makes it difficult to assign a percept to the functioning of a particular set of neurons or circuits. Tackling such questions requires interaction between psychophysics and physiology, and a good example is to be found in the paper by Davis and Driver on page 791 of this issue¹. The particular issue of interest here is how the brain 'sees' so-called Kanizsa or subjective figures.

Following on the pioneering discovery by Hubel and Wiesel² that neurons in the mammalian visual cortex respond preferentially, often exclusively, to contours

of a narrow range of orientations in a given retinal location, the mapping of receptive fields has been extensively employed as a tool by visual neurophysiologists. But after working with the more immediate defining criteria — motion, colour, contrast, disparity, pattern shape and size — it has still not become clear how percepts are developed; receptive fields do not map in a simple, unambiguous way onto seen objects. The best one can do is to relate receptive fields to attributes of visual percepts, such as position, depth and colour.

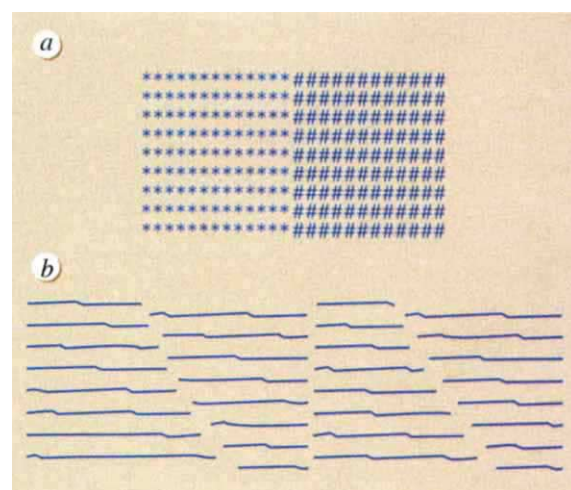
The notion of a border frequently arises in discussions of visual perception, and here the connection with orientation-selective receptive fields becomes obvious. But borders do not always have to be explicit. In 1900, the German psychologist Schumann drew attention to *Scheinkanten*, that is, situations in which one sees a border between two fields where in fact there is no explicitly drawn demarcation between them. This concept has gained prominence by the work of the Italian psychologist Kanizsa, who referred to them as *margini quasi-percettivi*. We now use the term illusory or subjective contours³. They appear as separators between a variety of textures (a in the figure), can be sketched in by distant inducing patterns (see Fig. 1b of Davis and Driver's paper), and share most of the properties of continuous borders or edges: they can be straight or curved, have the same

sensitivity to orientation differences⁴, and can participate in the induction of other illusions such as tilt⁵ (b here) and even three-dimensional effects⁶.

Can orientation-selective units in the visual cortex, originally mapped out by explicit lines and borders, also respond to illusory contours? Indeed they can, as was demonstrated by von der Heydt and Peterhans⁷, who even encountered the occasional cell in the visual cortex that responded better to an illusory than to a real contour.

The proposition that illusory contours enter the edge- or border-processing apparatus early on is strengthened by the findings of Davis and Driver, who employed an often-used approach to distinguish between 'early' and 'late' stages of visual processing. These terms need not be defined too rigorously, but the underlying principle is that some percepts are reasonably primitive and arise pretty soon after the retinal signals enter the cortex, whereas others are the result of more complex, interactive processing. For example, among many jumbled white letters, a single red one 'pops out' — that is, is recognized instantly — but a single L among many Ts needs some time to be detected. This procedure can be quantified by measuring the time it takes for a single odd feature to be detected as standing out among many background features. 'Early' processing is characterized by a reaction time that is fast and essentially independent of the number of background elements.

Using this approach Davis and Driver conclude that figures outlined by illusory contours do indeed fit into the 'early' category and justify the postulate that



Examples of illusory or subjective contours. a, An illusory edge appears as a separator between two fields of different texture. b, The sloping borders created by offsets in horizontal lines can participate in visual illusions — in this case the tilt effect, where the middle, strictly vertical illusory contour has a tilt induced in it by the surrounding tilted illusory contours. Another example (Fig. 1b, page 791) is where a triangle appears as an illusory figure induced by 'pacmen' (gaps in disks).

1. Davis, G. & Driver, J. *Nature* **371**, 791–793 (1994).
2. Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **148**, 574–591 (1959).
3. Petry, S. & Meyer, G. E. (eds) *The Perception of Illusory Contours* (Springer, New York, 1987).
4. Vogels, R. & Orban, G. A. *Vision Res.* **27**, 453–467 (1987).
5. Westheimer, G. *Vision Res.* **30**, 1913–1921 (1990).
6. Carman, G. J. & Welch, L. *Nature* **360**, 585–587 (1992).
7. von der Heydt, R. & Peterhans, E. *J. Neurosci.* **9**, 1731–1748 (1989).

they are processed by neuronal circuits in the visual cortex, something that is also pointed to by the equivalence of illusory and real contours in so many psychophysical and perceptual phenomena. But there are obvious differences in potency among illusory contours, which can be measured by their efficacy to participate in the tilt effect⁵, and this may yet help further to sharpen the concept of 'border' and its neural substrate.

We have here an example of convergence of neurophysiological and psycho-

physical approaches, where the understanding of a given phenomenon, in this case illusory contour, is enhanced when evidence obtained by one discipline is used in the design and interpretations of experiments in the other. Brain and behaviour are linked; researchers in the area are beginning to profit from their interdependence. □

Gerald Westheimer is in the Division of Neurobiology, University of California at Berkeley, Berkeley, California 94720, USA.

SUPERCONDUCTIVITY

Nomadic vortices

David Bishop

THE early euphoria associated with the discovery of the high- T_c superconductors has now been replaced by the sober realization that practical applications can only come from a vast amount of hard work, with success by no means assured. This is in large measure due to the pathological and mysterious behaviour of the magnetic vortices in these systems. On page 777 of this issue, Yao *et al.*¹ demonstrate an elegant way of helping to penetrate the mystery.

When a uniform magnetic field is applied to a type II superconductor, the field penetrates into the sample in quantized bundles of magnetic flux called magnetic vortices or magnetic flux lines. Microscopically these resemble whirlpools of circulating current. Any additional cur-

rent applied to the sample will induce a force on the vortices. If the vortices are free to move in response to this induced force they will dissipate energy and the superconducting state will be destroyed.

For almost every imagined technological application of superconductivity, the vortices must be pinned or prevented from moving if one is to obtain the benefits of using a superconductor. How to do this is the crux of the problem in applying superconductivity to create a new technology. There will be a huge economic payoff to those who can develop an optimal strategy to achieve this (and scientific papers for those who can't).

The first step in pinning the flux lines is to know something about their geometry: corrals for snakes or elephants need to be

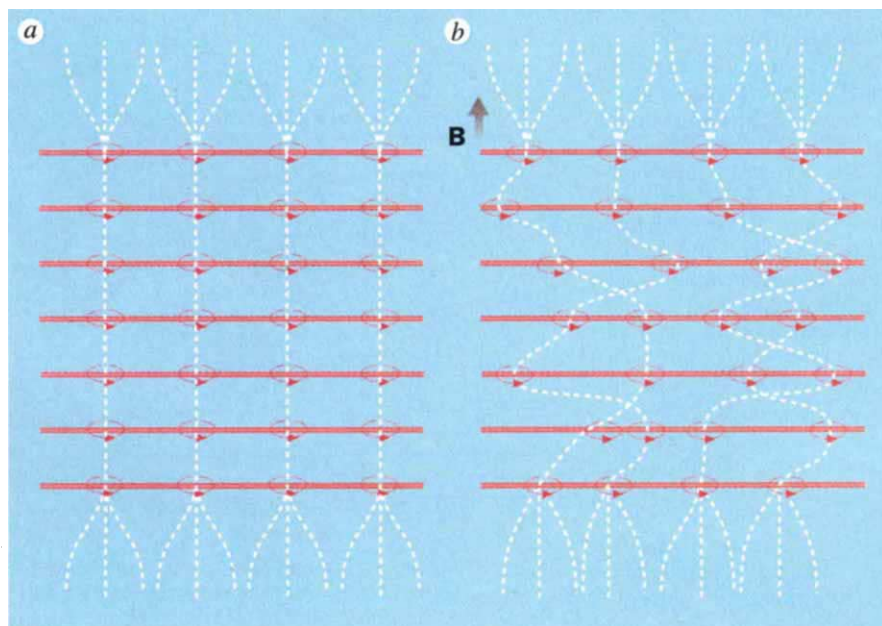
quite different objects. In the superconducting problem, it comes down to the issue of the dimensionality of the flux lines — are they three-dimensional lines or two-dimensional pancakes? The figure shows the two different types of vortices schematically. Pinning lines will prove to be much easier as one can grab them anywhere along their length to keep them from moving, whereas pancakes will each need to be individually pinned, a much harder task. The experiment by Yao *et al.*¹ contrives to get at the dimensionality in a most ingenious way.

The authors have used an old technique² for imaging flux lines — magnetic decoration — and extended it in a new direction. It is reminiscent of the experiment we all did as children when we used magnetic particles to visualize the field lines around a magnet: microscopic iron particles are used to decorate the location on the surface of the superconductor where the magnetic flux lines leave the sample. This is done by cooling a sample to low temperatures in an applied magnetic field, using an apparatus which allows one to evaporate an iron film onto the sample *in situ*. By having a small amount of helium gas in the can during the evaporation, instead of producing an iron film, one produces iron 'smoke' consisting of many particles of roughly 50 Å. These particles drift down onto the surface of the sample following the field lines of the magnetic vortices — a microscopic version of a homing beacon at an airport.

The particles thus strike the superconductor surface in locations where the flux lines leave, and they are held there by an 'atomic glue', the van der Waals forces. One can warm the sample up and use an electron microscope to locate the decorated positions. The magnetic forces are long-ranged but weak whereas the van der Waals forces are strong but short-ranged (so the particles stick well to the surface). This happy combination of circumstances is quite fortuitous and is what allows the technique to work as well as it does.

The method was first applied to magnetic domains by Bitter³ and used in superconductors in the pioneering work of Trauble and Essman⁴, and Sarma⁵. A number of groups have profitably used it to study flux lines and flux lattices in the high- T_c superconductors⁶, and to measure the flux quantum, pinning at grain boundaries, flux liquids, hexatics and vortex chains, and finally to measure the correlations of the vortices in the plane of the sample surface. But although the technique is easy to implement and fairly powerful, until now it has been limited to giving us information about a single slice of the flux lattice as it leaves the sample.

What Yao *et al.* have done will leave many researchers wondering, "Why didn't I think of that?" They have extended the technique to give information



a, Magnetic field lines (dashed lines) as they might appear on passing through a superconductor containing 'line' vortices — neatly correlated from layer to layer of copper oxide (red), so that the patterns of vortices seen on the top and bottom of the superconducting sample are perfectly correlated. b, What happens for 'pancake' vortices. The patterns on the top and bottom are entirely uncorrelated.

out of the plane of the surface by simultaneously decorating both sides of the sample and looking at correlations between the patterns on the top and bottom.

To see how this works, think about the limit in which the vortices are rigid, straight lines always parallel to the applied field. In that case, the patterns observed on the top and bottom would be completely correlated while in the opposite limit, that of two-dimensional pancake vortices, the patterns would be completely independent. As the authors reasoned, one can turn this around and use the correlations between the patterns seen on the top and bottom of the sample as a measure of the effective vortex dimensionality. Although it is simple in concept, in practice a number of issues such as sample heating during the evaporation required elegant experimental solutions which the authors devised en route.

Using their technique, Yao *et al.* were able to show that in single crystals of Bi:2212, at least in the regimes of field and temperature where they worked, the magnetic vortices are indeed line-like objects that extend throughout the entire sample. This result provides a natural explanation for the observation⁷ that line defects in this compound (such as ion tracks) provide much better pinning than point defects (such as vacancies or interstitial atoms). If the 'flux line' is indeed a line, then pinning it along its entire length in a line defect should be better than pinning it at one place. If they were points or two-dimensional pancake vortices then there should not have been much difference in the efficacy of point and line defects as pinning centres.

The problem of flux-line pinning in the oxide superconductors is one of the most important intellectual challenges in condensed matter physics today. It is impossible to solve this problem without microscopic structural information about the configuration of flux lines; after all, if asked to build a container for water, one's first question should be whether it is liquid, solid or gas. Questions of how much the vortex lines wander as they traverse the specimen have been hard to answer. The experiment by Yao *et al.* gives us a fine new weapon to add to the arsenal of experimental techniques available to attack the problem. □

David Bishop is in AT&T Bell Laboratories, 600 Mountain Avenue, PO Box 636, Murray Hill, New Jersey 07974, USA.

Turned on by insulin

Christopher G. Proud

ONE of the earliest effects of insulin and growth factors on animal cells is the stimulation of protein synthesis. The mechanisms involved remain obscure, despite the identification of a plethora of potential regulatory proteins¹. But progress is being made, the latest evidence coming in the form of papers by Pause *et al.* (page 762 of this issue²), Haystead *et al.* (*Journal of Biological Chemistry*³) and Lin *et al.* (*Science*⁴). We now have not

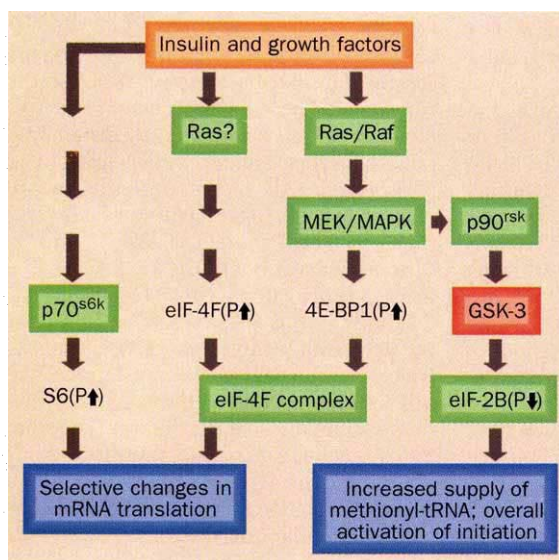
with only a few ribosomes. Following stimulation, these mRNAs acquire ribosomes and therefore shift into polyribosomes. There are at least two main classes of them. Some possess a sequence of about 8–14 pyrimidines at their extreme 5' termini⁵. These include mRNAs for ribosomal proteins and probably also translation factors. The 5' untranslated regions (5' UTRs) of the mRNAs in the second group are rich in self-complementary regions (secondary structure).

The potential significance of this becomes clear when we consider how eukaryotic ribosomes are generally thought to locate the AUG start codon — a major event in the initiation of translation, which is mediated by proteins termed eukaryotic initiation factors (eIFs).

All eukaryotic cytoplasmic mRNAs of cellular origin possess a so-called 'cap', containing a 7-methylguanosine moiety, at their 5' end. The initiation factor eIF-4E binds to the cap, and the ribosome is thought to interact initially with the mRNA at or close to the 5' end, through its interaction with eIF-4E and/or associated proteins. The ribosome (or rather its smaller 40S subunit) is then believed to move along the mRNA scanning for the first suitable AUG codon, from which translation commences. eIF-4E is part of a complex termed eIF-4F, which also contains eIF-4γ (of unknown but essential function) and eIF-4A, an RNA helicase; this helicase

apparently unwinds regions of secondary structure that hinder scanning⁶.

The newly identified proteins were found using a screen employing 'far-western' blotting. Pause *et al.*² screened an expression library of human complementary DNA for proteins which interact with eIF-4E. The positive plaques contained clones encoding two related proteins termed 4E-BP1 and 4E-BP2 (for eIF-4E binding proteins 1 and 2), of 118 and 120 amino acids respectively. Using several complementary approaches, the authors show that 4E-BP1 and 4E-BP2 each interact with eIF-4E and that they both *in vitro* and when overexpressed in human cells.



Signalling pathways involved in the control of translation initiation. p70^{s6k} lies on a distinct signalling pathway, whereas it seems that both 4E-BP1 and GSK-3/eIF-2B are regulated via the MAP kinase pathway. The protein kinases Raf and MEK lie upstream of MAP kinase, while p90^{rsk}, which is activated by MAP kinase, can phosphorylate and inactivate GSK-3. eIF-4F is subject to phosphorylation on two of its three subunits (eIF-4E and eIF-4γ) at multiple sites, but the signalling pathways involved (as well as the functional significance of these modifications) are unclear. Components in green boxes are activated by insulin or growth factors, whereas GSK-3 (red box) is inactivated. (P↑) and (P↓) indicate increases and decreases in phosphorylation for selected proteins.

only the identification of hitherto unknown players which modulate the translation of messenger RNA, but also strong indications as to a link between at least one of these proteins and a well-characterized pathway turned on by insulin and by growth factors. These and other recent findings mean that a clearer picture of how translation is controlled is at last beginning to emerge.

The rapid stimulation of protein synthesis is manifested both in a general rise in rates of mRNA translation and in pronounced increases in the translation of certain mRNAs. This second, selective phenomenon affects mRNAs which, before stimulation of the cells, are rather poorly translated and are often associated

1. Yao, Z., Yoon, S., Dai, H., Fan, S. & Lieber, C. M. *Nature* **371**, 777–779 (1994).

2. Bishop, D. J. *et al.* *Science* **255**, 165–172 (1992).

3. Bitter, F. *Phys. Rev.* **38**, 1903 (1931).

4. Trauble, H. & Essman, U. J. *appl. Phys.* **39**, 4052 (1968).

5. Sarma, N. V. *Phil. Mag.* **17**, 1233 (1968).

6. Bishop, D. J. in *The Vortex State* (eds Bontemps, N. *et al.*) 99–123 (Kluwer, Dordrecht, 1994).

7. Nelson, D. R. in *Phase Transitions and Relaxation in Systems with Competing Energy Scales* (eds Riste, T. & Sherrington, D.) 95–117 (Kluwer, Dordrecht, 1993).

Seen in the context of the cloning of a protein termed PHAS-I earlier this year⁷, the potential significance of the two 'new' proteins became clear. PHAS-I was first identified as a protein whose phosphorylation was rapidly enhanced by insulin in rat fat cells⁸, and 4E-BP1 is its human counterpart. Pause *et al.*² and Lin *et al.*⁴ went on to demonstrate that insulin treatment of fat cells disrupts the interaction of 4E-BP1 with eIF-4E (presumably as a consequence of the phosphorylation of 4E-BP1). This interaction seems to inhibit translation, so these observations provide evidence of a mechanism by which insulin (and other agents, such as certain growth factors, which also increase the phosphorylation of this or a closely related protein⁹) could stimulate translation. The protein seems to be present in all insulin-sensitive tissues^{4,7,10,11}.

It is not yet clear how the interaction with 4E-BP1 blocks translation, but it probably impairs the binding of eIF-4E to eIF-4A or eIF-4γ (or both). The working hypothesis is therefore as follows. Enhanced phosphorylation of 4E-BP1 leads to its dissociation from eIF-4E, allowing this protein to interact with eIF-4A/4γ to form a functional complex. This would particularly facilitate the translation of mRNAs whose 5'UTRs contain significant secondary structure and are therefore especially dependent on eIF-4A function. An example may be the mRNA for ornithine decarboxylase: its 5'UTR is highly structured and its translation is increased by insulin¹² or by overexpression of eIF-4E (ref. 13).

Which protein kinase (or kinases) mediates the increase in 4E-BP1 phosphorylation in response to insulin? In their paper Haystead *et al.*³ show that MAP kinase phosphorylates the rat homologue of 4E-BP1 (PHAS-I) at a single serine residue (Ser 64), which is also phosphorylated in response to insulin in adipocytes. MAP kinase is activated by insulin and growth factors in many types of cell. Indeed, Lin *et al.* show that the two isoforms of MAP kinase are the major PHAS-I kinases in epidermal growth factor-stimulated 3T3-L1 adipocytes⁴. Furthermore, phosphorylation of PHAS-I *in vitro* by MAP kinase blocks its ability to bind to eIF-4E⁴. But the protein contains phosphothreonine as well as phosphoserine, so it can also be phosphorylated by at least one other protein kinase. It is not yet clear whether 4E-BP2 is also phosphorylated or, if it is, under what conditions this might be brought about.

The picture can be broadened to include other notable developments. It has been known for many years that the ribosomal protein S6, in the 40S subunit, undergoes increased phosphorylation in response to many agents that stimulate translation, through the activation of a specific protein kinase (p70^{s6k}). Despite

recent detailed studies, the mechanism of activation of p70^{s6k} remains obscure¹⁴. The function of S6 phosphorylation had also been unclear, but work published in May¹⁵ has shed fresh light on this question. Rapamycin is an immunosuppressant that blocks activation of p70^{s6k} and S6 phosphorylation, and Jefferies *et al.*¹⁵ showed that it also blocks the recruitment of mRNAs containing polypyrimidine tracts into polysomes; this indicates that phosphorylation of S6 is central to this process.

Another potential regulatory mechanism which has recently been revealed involves the initiation factor eIF-2B, a guanine-nucleotide-exchange factor which maintains the initiator transferRNA-binding factor, eIF-2, in its active GTP-bound state, thus ensuring the supply of initiator methionyl-tRNA to the ribosome. The exchange activity of this factor is acutely stimulated by insulin and growth factors, and has been shown to be phosphorylated and apparently inactivated by glycogen-synthase kinase-3 (GSK-3) (ref. 16 and G. I. Welsh, unpublished observations). Because GSK-3 is itself inactivated in response to insulin^{16,17}, this provides a pathway for the activation of initiator tRNA binding, which is essential for the translation of all mRNAs and should lead to an across-the-board stimulation of translation. Together these mechanisms could therefore bring about both the general activation of initiation and selective increases in the expression of particular mRNAs. □

Christopher G. Proud is in the Department of Biochemistry, School of Medical Sciences, University of Bristol, University Walk, Bristol BS8 1TD, UK.

1. Redpath, N. T. & Proud, C. G. *Biochim. biophys. Acta* **1220**, 147–162 (1994).
2. Pause, A., Belsham, G. J., Gingras, A.-C., Donzé, O., Lin, T.-A., Lawrence, J. C. Jr & Sonenberg, N. *et al. Nature* **371**, 762–767 (1994).
3. Haystead, T. A. J., Haystead, C. M. M., Hu, C., Lin, T.-A. & Lawrence, J. C. Jr *J. Biol. Chem.* **269**, 23185–23191 (1994).
4. Lin, T.-A., Kong, X., Haystead, T. A. J., Pause, A., Belsham, G., Sonenberg, N. & Lawrence, J. C. Jr *Science* **266**, 635–656 (1994).
5. Morris, D. R., Kakegawa, T., Kaspar, R. L. & White, M. W. *Biochemistry* **32**, 2931–2937 (1993).
6. Sonenberg, N. *Gene Expression* **3**, 317–323 (1993).
7. Hu, C., Pang, S., Kong, X., Vellica, M. & Lawrence, J. C. *Proc. natn. Acad. Sci. U.S.A.* **91**, 3730–3734 (1994).
8. Belsham, G. J. & Denton, R. M. *Biochem. Soc. Trans.* **8**, 382–383 (1980).
9. Blackshear, P. J., Nemenoff, R. A. & Avruch, J. *Biochem. J.* **214**, 11–19 (1982).
10. Diggle, T. A. & Denton, R. M. *Biochem. J.* **282**, 729–736 (1992).
11. Diggle, T. A., Bloomberg, G. B. & Denton, R. M. *Biochem. J.* (in the press).
12. Manziella, J. M., Rychlik, W., Rhoads, R. E., Hershey, J. W. & Blackshear, P. J. *J. Biol. Chem.* **266**, 2383–2389 (1991).
13. Shantz, L. M. & Pegg, A. E. *Cancer Res.* **54**, 2313–2316 (1994).
14. Ming, X. F. *et al. Nature* **371**, 426–429 (1994).
15. Jefferies, H. B. J., Reinhard, G., Kozma, S. C. & Thomas, G. *Proc. natn. Acad. Sci. U.S.A.* **91**, 4441–4445 (1994).
16. Welsh, G. I. & Proud, C. G. *Biochem. J.* **294**, 625–629 (1993).
17. Hughes, K., Ramakrishna, S., Benjamin, W. B. & Woodgett, J. R. *Biochem. J.* **288**, 309–314 (1992).

Area of no return

FRICION is rather a subtle phenomenon. On the atomic scale, all surfaces are very rough. Two objects in contact touch only at a few high points, which carry all the load. The high local pressure deforms them against each other, and they cold-weld together. The friction of sliding is the force needed to shear these microscopic welds.

Daedalus is now complicating this simple picture. He is creating surfaces slippery in one direction, but sticky in the other. Samples for electron microscopy are often 'shadowed' with metal. A beam of metal atoms, sprayed *in vacuo* from a low angle, paints each feature with a dramatic shadow of deposited metal on one side, like evening sunlight on a mountainous landscape. So Daedalus is shadowing engineering surfaces with solid lubricant molecules or weld-weakening agents. They coat every microscopic peak on the surface, but on one side only. A surface so coated should show anisotropic friction.

Push it against another surface, and the two will weld together at their points of contact. But these microwelds will all be weakened on one side by the asymmetric deposits of coating agent. Slide the joint in one direction, and the weak sides will be put under tension: they will readily fail by crack propagation. Slide the joint the opposite way, and the tension will come on the strong sides of the welds. They will resist the push.

The best materials for the trials are copper and aluminium, whose surfaces normally seize together readily. Properly shadowed, however, they should be remarkably slippery in the 'easy' direction. When the technique has been perfected, Daedalus's one-way sliding surfaces will revolutionize the bolting and fastening industry.

One-way nuts will easily spin down their bolts, but will bind hard if turned the other way. So simple vibration will automatically ratchet them tight. One-way nails, wedges and rivets will similarly glide smoothly into place, and will then grip with unyielding tenacity. Complete one-way self-assembling mechanisms may even be possible. Vibrate the system, and all the joints will slide along their rods, unfolding the whole structure firmly and irreversibly. Even the most violent buffeting or shaking will not imperil it, but will merely tighten its joints more firmly. Production engineers will be delighted; but scrap-merchants and repair-men, unable to dismantle the new products, will grind their teeth in frustration. Daedalus's one-way components will be even more unremovable than those infuriating crosshead screws.

David Jones

Anarchy in the beehive

SIR — Worker bees in eusocial Hymenoptera colonies often have functional ovaries which can produce fertile male eggs¹. The sons of a worker's 'super sister' are only $\frac{3}{4}$ as much related to her as her own sons, and only $\frac{1}{2}$ as much related to her as sons of the queen. This leads to the possibility of intense competition among workers for the production of

(3) workers carrying queen-laid drone eggs through the queen excluder.

A candidate colony was reported from Ipswich, Queensland in late winter, before swarming season. The colony was normal in every respect except that most drone-sized cells (>100) above the queen excluder contained brood. No worker brood was present above the queen

One additional drone was descended from a worker of another patriline marked by the 170 allele at locus A107. It is overwhelmingly likely that the remaining 48 drones were derived from workers of a single patriline because only one non-queen allele was found at each microsatellite locus in all drones examined.

It is almost impossible for our sample to have contained any drones laid by the queen. In 42 cases this possibility can be definitely excluded because these drones did not carry a queen allele at one or more loci. In the other six cases, it is possible but unlikely, as workers would have to have moved queen-laid eggs through the queen excluder⁷.

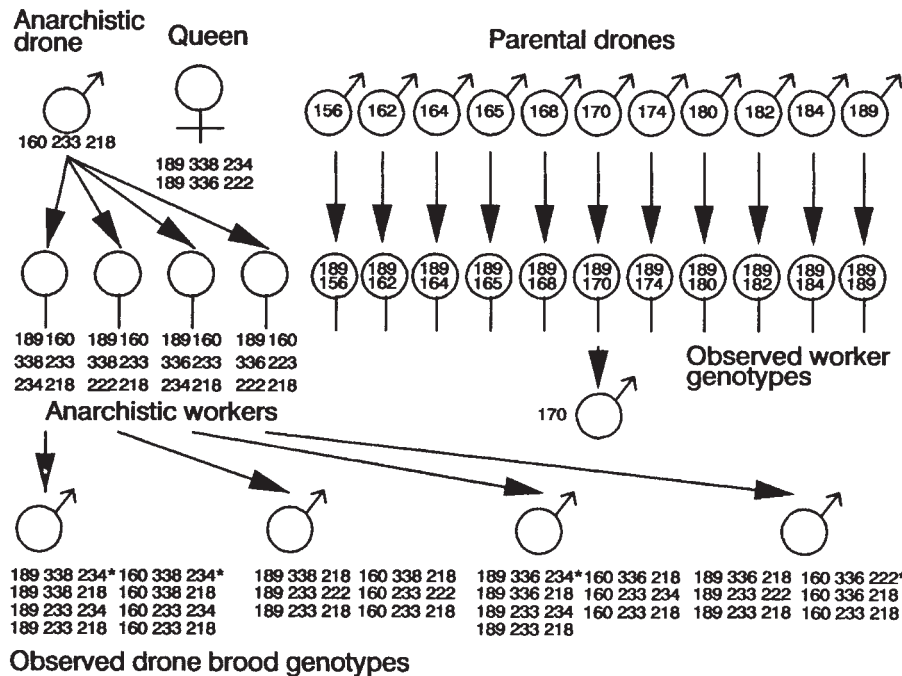
We conclude that anarchistic behaviour is genetically determined, that one of the males mating this queen passed an ability to his worker offspring to evade the policing mechanism. If queen-laid eggs are protected from policing by a queen-derived pheromone^{5,7}, then it is likely that several bees of this patriline were able both to develop ovaries and to produce an acceptably queen-like pheromone that protected their eggs. It seems likely that workers of this patriline were more queen-like⁹ than normal bees. There is large genetic variance in the speed at which individuals develop ovaries and the probability that they will become 'false queens' in queenless colonies⁴.

This anarchistic behaviour could rapidly spread in honey bee populations. The fitness of these bees is higher than that of workers that cannot evade policing. Selection on workers to police this behaviour is expected to be strong given the number of times the queen mates². These results support the worker policing hypothesis, while demonstrating a previously unrecognized dynamic equilibrium between reproductive conflict and cooperation.

Benjamin P. Oldroyd
Adam J. Smolenski
Jean-Marie Cornuet
Ross H. Crozler
School of Genetics and Human Variation, La Trobe University, Bundoora, Victoria 3083, Australia

J.-M. C. is on sabbatical from Laboratoire de Neurobiologie Comparée des Invertébrés, INRA CNRS URA 11290, La Guyonnerie, BP 23, 91440 Bures-sur-Yvette, France.

- Oster, G. F. & Wilson, E. O. *Caste and Ecology in the Social Insects* (Princeton University Press, 1978).
- Ratnieks, F. L. W. *Am. Nat.* **132**, 217–236 (1988).
- Visscher, P. K. *Behav. Ecol. Sociobiol.* **25**, 247–254 (1989).
- Page, R. E. & Erickson, E. H. *Behav. Ecol. Sociobiol.* **23**, 117–126 (1988).
- Ratnieks, F. L. W. & Visscher, P. K. *Nature* **342**, 796–797 (1989).
- Starr, C. K. in *Sperm Competition and the Evolution of Animal Mating Systems* (ed. Smith, R. L.) 428–459 (Academic, Orlando, 1984).
- Ratnieks, F. L. W. *Behav. Ecol. Sociobiol.* **32**, 191–198 (1993).
- Estoup, A., Solignac, M. & Cornuet, J.-M. *Proc. R. Soc.* **258**, 1–7 (1994).
- Moritz, R. F. A. & Hillesheim, E. *Behav. Ecol. Sociobiol.* **17**, 87–98 (1985).



Pedigree of a honeybee colony determined from three microsatellite loci. Drones marked with an asterisk are unique to that family, and indicate that at least three workers laid eggs.

males². In many species, however, worker reproduction is extremely rare^{3,4}, probably because it is suppressed by a mutual 'policing' behaviour in which workers prevent others from laying eggs^{5,6}. In honeybees there is good evidence that the mechanism of worker policing involves the marking of queen-laid eggs by a pheromone. In queen-right colonies, workers can identify worker-laid eggs by the lack of this putative pheromone, and eat them⁷. This system works very well most of the time, but is open to the invasion of 'anarchistic' behaviour. Here we report a colony where workers of a single patriline 'cheated' to produce many of the colony's male offspring.

Occasionally colonies are seen where most drones are likely to be offspring of workers, not the queen⁴. We asked beekeepers to report colonies where drone pupae were present above a queen excluder in otherwise normal colonies. Because a queen cannot pass through a queen excluder, drone pupae above it can only result from: (1) a second queen above the queen excluder; (2) laying worker(s); or

excluder. We obtained samples of worker brood from below the queen excluder, and adult workers and 49 drone pupae from above the excluder. Drone pupae were only present in drone-sized cells. We extracted DNA from these samples and analysed it for variation at three microsatellite loci⁸ to determine the paternity and maternity of each bee. Analysis of 128 adult and 74 pupal workers with microsatellite A107 revealed that every worker carried a 189 base-pair (bp) allele. Therefore the single queen heading the colony was homozygous for the 189 bp allele. Twelve A107 alleles were present in the worker progeny indicating that the queen had mated with at least 12 drones. Examination of 12 workers with additional microsatellites A14 and A76 revealed that the queen was heterozygous 338/336 and 234/222 for these loci, as all workers carried at least one of these alleles.

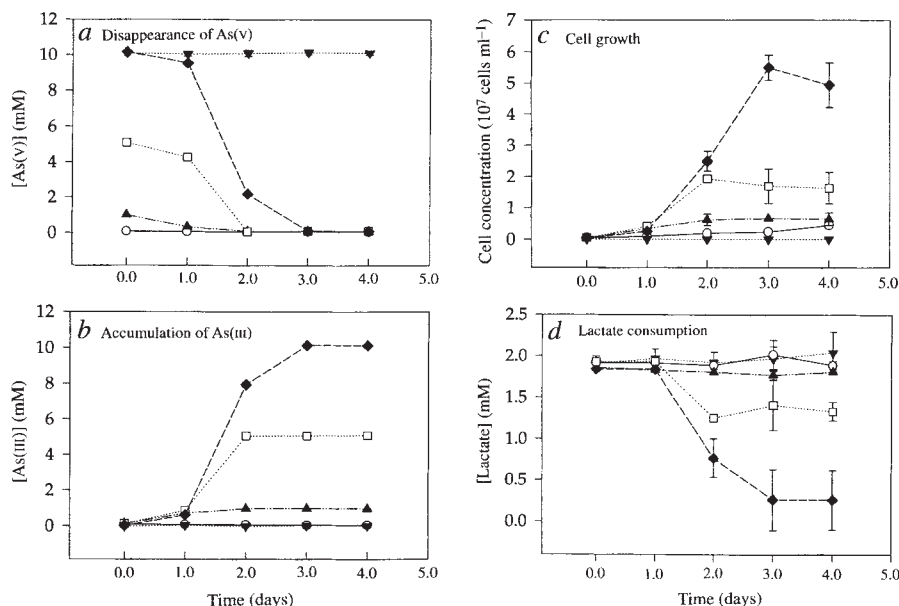
From the distribution of microsatellite alleles among the drone pupae, the most parsimonious hypothesis is that they are offspring of at least three workers sired by the same drone of genotype 160-233-218.

Microbe grows by reducing arsenic

STR — Arsenic is a legendary poison, renowned in its effectiveness against heroines, gentlemen, rats and wood rot. Arsenic is lethal to microorganisms, yet certain bacteria and phytoplankton are known to survive arsenic exposure and transform arsenic species through oxidation¹, reduction² and methylation³; to date, these transformations are thought to occur either incidentally or through specific detoxification mechanisms⁴. There is an extraordinary diversity in microbial use of inorganic compounds as terminal electron acceptors in anaerobic respiration, from the common Fe(III) (ref. 5) and Mn(IV) (ref. 6) to the exotic U(VI) (ref. 7). Thermodynamic calculations show that dissimilatory arsenate reduction could also provide sufficient energy to sustain microbial growth, although the universal toxicity of arsenic militates against such an idea. We report here, however, the discovery and isolation of a microorganism, designated MIT-13, that reduces As(V) to As(III) and gains energy for growth from this reduction.

The isolate MIT-13 was obtained from arsenic-contaminated sediments of the Aberjona watershed in eastern Massachusetts. Anoxic sediments were inoculated into minimal medium to develop an enrichment culture, and a pure culture was obtained by successive colony isolations on solid medium.

To demonstrate the dependence of cell growth and substrate consumption on arsenate concentration, we inoculated MIT-13 into medium supplemented with 2 mM lactate and various concentrations of arsenate. Within 4 days, complete disappearance of As(V) (a in the figure) was accompanied by stoichiometric appearance of As(III) (b). Cells grew as As(V) was reduced (doubling time approximately 14 h with 10 mM arsenate) and reached final numbers that were proportional to initial As(V) concentrations (c). Lactate disappearance was observed in proportion to As(III) production (d). Uninoculated cultures showed no arsenic reduction, lactate consumption, or cell growth. Inoculated cultures deprived of As(V) showed minimal cell growth and lactate consumption. Addition of formaldehyde to a final concentration of 4% completely inhibited both cell growth and arsenate reduction.



a, As(V) concentration; b, As(III) concentration; c, cell growth; d, lactate consumption over a 4-day incubation of MIT-13 in anaerobic medium supplemented with lactate as substrate and As(V) as electron acceptor. Cells were inoculated into duplicate cultures of arsenate-lactate medium containing 0 (○), 1 (▲), 5 (□), and 10 mM (◆) arsenate. An additional set of bottles with 10 mM arsenate was prepared, not inoculated, and treated exactly as the others (▼). At 24-h intervals, samples were taken from each bottle and analysed for cell concentration, lactate concentration, and arsenic concentration and speciation. Cell concentrations were determined by DAPI staining and direct counting with epifluorescence microscopy. Lactate concentrations were measured by HPLC with separation on a polysulphonate ion exclusion column (Hamilton PRP×300) and detection by ultraviolet absorbance at 210 nm. Arsenic concentrations and speciation were measured by the molybdenum blue spectrophotometric assay⁹. Medium for enrichment and purification consisted of freshwater minimal salts medium buffered at pH 6.8 with NaHCO₃ and supplemented with 0.1 mM cysteine, 2 mM lactate and 10 mM sodium arsenate. Medium was reduced with PdCl₂ (0.005%) + H₂, and immediately after inoculation headspaces were changed to a N₂:CO₂ ratio of 80:20. All incubations were carried out under N₂:CO₂:80:20 atmosphere at 22 °C in darkness.

Microscopic examination of MIT-13 reveals a vibrio-shaped, highly motile microorganism of approximately 1 μm in length and 0.1 μm in width. Cells are unable to use acetate as a substrate, but using lactate they are able to grow on sulphate as well as arsenate. Arsenate reduction and cell growth are transiently inhibited by molybdate (50% increase in lag phase at 20 mM) but not by sulphate. Complete reduction of 10 mM arsenate in the presence of 2 mM lactate and 20 mM sulphate demonstrates that As(V) is the preferred electron acceptor in cultures previously grown on As(V). Oxygen completely inhibits both arsenate reduction and cell growth.

Before the isolation of MIT-13, we incubated sediment samples from the Aberjona watershed in anoxic freshwater medium supplemented with typical fermentation end-products (acetate, lactate and butyrate) and solid iron (III) arsenate, a compound representative of sedimentary arsenic⁸. The enrichment cultures did indeed cause dissolution and reduction of the arsenate to arsenite within several days. Autoclaved and formaldehyde-killed cultures did not dissolve nor reduce the ferric arsenate, showing that active metabolism is necessary to transform the arsenate (data not shown).

Our evidence shows that rapid, extensive arsenate reduction and dissolution can be accomplished by microbial activity, both in pure culture and in natural consortia, and suggests that microbial metabolism plays a greater, more direct role in arsenic cycling than previously suspected.

Dianne Ahmann

Department of Civil and Environmental Engineering,

Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

A. Lynn Roberts

Department of Geography and Environmental Engineering,

Johns Hopkins University, Baltimore, Maryland 21218, USA

Lee R. Krumholz

Department of Botany and Microbiology, University of Oklahoma, Norman, Oklahoma 73019, USA

François M. M. Morel

Department of Geology, Princeton University, Princeton, New Jersey 08540, USA

- Osborne, F. H. & Ehrlich, H. E. *J. appl. Bact.* **41**, 295–305 (1976).
- Sanders, J. G. & Windom, H. L. *Est. Coast. mar. Sci.* **10**, 555–567 (1980).
- Anderson, L. C. & Bruland, K. W. *Environ. Sci. Tech.* **25**, 420–426 (1991).
- Sliver, S. & Keach, D. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6114–6118 (1982).
- Lovley, D. R. & Phillips, E. J. *J. Appl. Env. Microbiol.* **54**, 1472–1480 (1988).
- Myers, C. R. & Neilson, K. H. *Science* **240**, 1319–1321 (1988).
- Lovley, D. R. *et al. Nature* **350**, 413–416 (1991).
- Ferguson, J. F. & Gavis, J. J. *Water Res.* **6**, 1259–1274 (1972).
- Johnson, D. L. & Pilson, M. E. *Analyt. chim. Acta* **58**, 289–299 (1972).

How old is 'Boxgrove man'?

SIR — Roberts *et al.*¹ concluded that the human tibia discovered at Amey's Eartham pit, near Boxgrove, Sussex, UK, is between 524,000 and 478,000 years old (oxygen isotope stage 13), older than the Anglian ice-age of stage 12 (ref. 1). This conclusion is based on 'correlative mammalian biostratigraphy' and comparative geomorphology. But the comparison of mammalian fossils from different sites is not underpinned by any independent geochronological age estimates determined by physical-chemical means. Nor is the geomorphological argument based on the fluvial gravels of an early River Thames at Little Oakley in East Anglia, UK unequivocal². This is not surprising because it is notoriously difficult to date individual events and their fossils in the

with oxygen isotope stratigraphy⁴. Ratios of 0.29 are correlated with oxygen isotope stage 11 (423,000–362,000 years before present⁵), and thus provide an approximate age of about 400,000 years for 'Boxgrove man' — an entire ice-age younger than the previous estimate¹.

Amplification is possible by comparison of marine and non-marine amino acid data, because the epimerization rate of *Littorina littorea* is similar to some terrestrial molluscs (gastropods and bivalves). At the interglacial site of Swanscombe in the lower Thames valley the mean ratio of the non-marine fossil molluscs is 0.3 ± 0.017 ($n = 34$). This is also correlated with oxygen isotope stage 11 by means of a similar, but non-marine, amino acid geochronological model⁶, that is constrained by stratigraphical and geomorphological age relationships of other deposits throughout southern Britain. The fossiliferous Swanscombe beds lie within the Orsett Heath gravel that directly overlies the Hornchurch Till (boulder clay) of the Anglian glaciation². Thus, the Swanscombe amino-acid ratios of 0.3 must be post-Anglian in age. Because the Anglian is correlated with oxygen isotope stage 12 (ref. 2) (478,000–423,000 years before present), it follows that the Swanscombe beds, and their equivalents at

GEOCHRONOLOGICAL AGE ESTIMATES FROM BOXGROVE

Thermoluminescence ¹²	175.3–229.4
Electron spin resonance ¹³	205–281
Uranium series ¹⁴	> 350
Aminostratigraphy (marine species) ^{8,9}	303–524
Aminostratigraphy (terrestrial species) ⁸	303–339

Age estimates are $\times 1,000$ years before present.

Littorina littorea nor the terrestrial gastropod *Tichia hispida*. Taking the range of means for these species⁸ together with the data given in Bowen and Sykes' table, and then using Bowen's correlation with the oxygen isotope record⁹, an age spread between stages 9 and 13 is indicated. Although the ratios from amino-acid epimerization at Boxgrove may be interpreted as internally consistent, this does not mean that they are valid when applied to wider aminostratigraphic age correlations. This point is well illustrated by ratios from Waverley Wood⁹ which are consistently greater than those from the Cromer Forest Bed, even though the latter site contains the ancestral form in the lineage of the water vole¹⁰ to that found at Waverley Wood.

Bowen and Sykes have still not addressed the chronologically important taxonomic differences between the Boxgrove mammalian fauna and the faunas from other sites, such as Barnham¹¹, that they have dated to OIS 11. There is consensus among many Quaternary researchers in Europe² that Boxgrove cannot be the same age as the Swanscombe interglacial stage and must therefore, as advocated in our Letter¹, date from the end of the Cromerian complex. (Details of references supporting our proposed chronology are available from M. B. R. on request.)

M. B. Roberts

*Institute of Archaeology,
University College,
London WC1H 0PY, UK*

AMINO-ACID RATIOS FROM BOXGROVE FOSSIL GASTROPODS

Lab No.*	Species	D-alloisoleucine/ L-isoleucine ratio	Stratigraphic unit
LOND 8	<i>Nucella lapillus</i>	0.265 ± 0.013 (5)	Top of Slindon Sands
LOND 501	<i>N. lapillus</i>	0.29 ± 0.025 (4)	Intra Slindon Sands
LOND 503	<i>L. saxatilis</i>	0.308 ± 0.032 (5)	Intra Slindon Sands
LOND 346	<i>L. saxatilis</i>	0.281 ± 0.009 (4)	Intra Slindon Sands
LOND 502	<i>N. lapillus</i>	0.298 ± 0.019 (9)	Basal boulder beach

*Refers to London University. The amino-acid geochronology laboratory is now in the Earth Sciences Department, University of Wales, Cardiff. Data are standardized to *L. littorea*³ and refer to the total hydrolysate.

the Middle Pleistocene of northwest Europe because of the uncertainties associated with applicable dating methods and paucity of suitable materials for dating.

Much progress has been made in recent years using ratios of the protein amino acids in fossils to estimate the age of molluscs. Living organisms contain only L-configuration amino acids. When they die, the L-configuration amino acids convert to those of the D-configuration (racemization). This is a reversible reaction and it results in increased D/L ratios in fossils through time until a D/L equilibrium ratio is reached: the older the fossil, the higher its D/L ratio will be. We applied this method to estimate the age of 27 separate fragments of marine gastropod fossils, collected for us by M. B. Roberts and M. R. Bates, from the Slindon sands and Basal Beach gravel which underlie the beds in which the human tibia was discovered. Using high-performance liquid chromatography, we measured the extent of the conversion of the protein amino acid L-isoleucine to its non-diastereoisomer D-alloisoleucine (epimerization).

The mean D-alloisoleucine/L-isoleucine ratio of all the samples is 0.29 ± 0.025 ($n = 27$) (see table). This can be converted into a calendar timescale using a model that correlates amino-acid ratios, calibrated by independent age estimates,

Boxgrove, are younger and are thus correlated with oxygen isotope stage 11 (423,000–362,000 years before present). For comparison, the standard non-marine amino acid ratio for oxygen isotope stage 13 is 0.035 ± 0.01 ($n = 9$) (ref. 6). In our view, there are too many uncertainties in the various dating methods to be able to assign a definitive date for the occurrence of early hominids in the British Isles.

D. Q. Bowen

G. A. Sykes

*Department of Earth Sciences,
University of Wales,
Cardiff CF1 3YE, UK*

ROBERTS REPLIES — Bowen and Sykes state that the pre-Anglian age proposed for the Boxgrove tibia is not underpinned by any geochronological age estimate determined by physical-chemical means. We applied various methods of age estimation to the sediments and the fossils contained therein at Boxgrove¹, but the fact is that the range of dates obtained was so varied that we used mammalian biostratigraphy, from a European database⁷, to construct a more accurate and testable chronological model when combined with detailed lithostratigraphic data.

Turning to the amino-acid data, Bowen and Sykes do not include in their table the two runs that were actually carried out on

1. Roberts, M.B. *et al.* *Nature* **369**, 311–313 (1994).
2. Bridgland, D.R. *Quaternary of the Thames*, (Joint Nature Conservation Committee, Chapman and Hall, London, 1994).
3. Bowen, D.Q. *et al.* *Quat. Sci. Rev.*, **4**, 279–318 (1985).
4. Bowen, D.Q. & Sykes, G.A. *Phil. Trans. R. Soc. Lond.* **B318**, 619–635 (1988).
5. Imbrie, J. *et al.* in *Milankovich and Climate Part 1* (eds Berger, A.L. *et al.*) 269–305 (Reidel, Dordrecht, 1984).
6. Bowen, D.Q. *et al.* *Nature* **340**, 49–51 (1989).
7. van Kolfschoten, T. *Medelingen Rijks Geologische Dienst* **43-3**, 1–69 (1990).
8. Bowen, D.Q. & Sykes, G.A. in *The Middle Pleistocene Site at Boxgrove, West Sussex, UK* (eds Roberts M. B. & Parfitt S. A.) (English Heritage Monograph Series, London, in the press).
9. Bowen, D.Q. in *Central East Anglia and the Fen Basin* (eds Lewis S. G. *et al.*) 21–24 (Quaternary Research Association, London, 1991).
10. Currant, A. P. in Shotton *et al.* *J. Quat. Sci.* **8**, 293–325 (1993).
11. Ashton, N. M. *et al.* *J. Geol. Soc.* **151**, 599–605 (1994).
12. Parks, D. A. & Rendall, H. M. *J. Quat. Sci.* **7**, 99–107 (1992).
13. Grün, R. in *The Middle Pleistocene Site at Boxgrove, West Sussex, UK* (eds Roberts M. B. & Parfitt S. A.) (English Heritage Monograph Series, London, in the press).
14. Rae, A. M. in *The Middle Pleistocene Site at Boxgrove, West Sussex, UK* (eds Roberts M. B. & Parfitt S. A.) (English Heritage Monograph Series, London, in the press).

Alu alert

SIR — The presence of Alu elements in protein coding regions has been a controversial issue for some time¹ and is periodically reactivated²⁻⁴.

The polymorphic short interspersed elements (approximately 300 base pairs long) of the Alu family are ubiquitous in the genome of primates. They constitute about 5 per cent of the human genome, but can also be found in dense clusters of up to two Alu elements per kilobase⁵. In relation to regular protein transcription units, Alu elements are often found in introns, or in 3' untranslated regions. A single base deletion or insertion while sequencing a genomic clone or a complementary DNA can easily cause Alu regions to be mistakenly incorporated into open reading frames¹. Also, cDNA libraries can contain partially spliced mRNA, be contaminated with genomic material, or contain rearranged clones from artefactual ligation and recombination.

In such rearranged cDNAs, Alu elements will probably be inserted within bona fide exon sequences. A few Alu elements are still transcriptionally active, and are thought to be spread within the genome through retrotransposition. This suggests a mechanism by which Alu elements might have played an important role in the recent evolution of certain primate proteins⁶; but the sequence data supporting such a hypothesis are difficult to distinguish from the many possible artefacts cited above⁴. A recent example of interpretational difficulties comes from BRCA1 (breast cancer) mRNA. One of the rare forms of this mRNA contains exon 4, composed entirely of an Alu cassette. As this was not mentioned by Miki *et al.*⁷ and exon 4 was found only in one of 15 variants of BRCA1 mRNA, this particular cDNA is probably a splicing error that results in a non-functional product.

While the issue is still in debate, an increasing number of Alu-containing protein sequences, conceptually translated from cDNA or genomic open reading frames, are finding their way into data bases (we detected at least 37 entries in NR, the NCBI non-redundant database). Despite the introduction of Alu warning entries in protein databases¹, the presence of Alu-like segments often escape detection by the original authors, thus generating entries with no mention of this feature. These potentially anomalous protein sequences, then, constitute a growing source of misleading homology matches.

To help researchers to detect Alu-like segments in putative coding regions early on, a select subset of 325 Alu elements is now available for a quick search via the NCBI BLAST server (blast@ncbi.nlm.nih.gov). This subset, its conceptual translation, and the various

programs of the BLAST suite allow protein-protein, nucleotide-protein and nucleotide-nucleotide comparisons. This small selection of Alu elements is highly representative and can detect 99 per cent of all known Alu elements in GenBank. In conjunction with the xblast⁸ program, search against this database can be used for reliably masking all Alu-like segments from large (for example, genomic contig) or multiple (partial cDNA) queries before scanning the protein databases or GenBank. This step greatly improves the biological relevance of the output by drastically reducing reported hits; and limiting them to functionally meaningful regions of the sequences.

The 325-Alu subset, its conceptual translation and a help file are available by anonymous FTP (ncbi/pub/jmc/alu). Also 8 Alu consensus sequences, representing each of the currently defined subfamilies, have been incorporated as warning entries into GenBank (accession numbers U14567-U14574).

Jean-Michel Claverie*

Wojciech Makalowski†

National Center for Biotechnology Information,

National Library of Medicine, NIH, Bethesda, Maryland 20894, USA

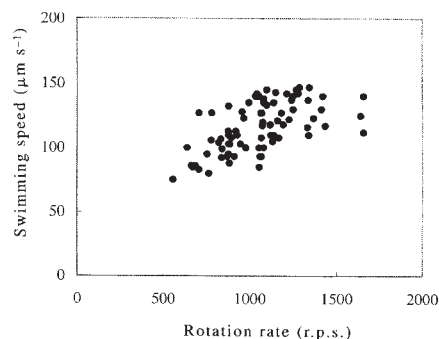
*(jmc@ncbi.nlm.nih.gov); †(makalow@ncbi.nlm.nih.gov)

1. Claverie, J.-M. *Genomics* **12**, 838-841 (1992).
2. Margalit, H. *et al. Cell* **78**, 173-174 (1994).
3. Tugendreich, S. *et al. Nature* **370**, 106 (1994).
4. Zietkiewicz, E. *et al. Science* **265**, 1110-1111 (1994).
5. Iris, F. J. M. *et al. Nature Genet.* **3**, 137-145 (1993).
6. Makalowski, W. *et al. Trends Genet.* **10**, 188 (1994).
7. Miki, Y. *et al. Science* **266**, 66-71 (1994).
8. Claverie, J.-M. in *Automated DNA Sequencing and Analysis* (eds Adams, M. *et al.*) 267 (Academic, New York, 1994).

Very fast flagellar rotation

SIR — Bacteria swim by rotating their helical flagellar filaments. Each filament is driven with a flagellar motor embedded in the cell surface. The rotation rates of the motor had been reported to be high, for example, 270 r.p.s. for *Escherichia coli*¹ and 170 r.p.s. for *Salmonella typhimurium*². We have observed remarkably faster rotation, up to 1,700 r.p.s. for polar flagellum of *Vibrio alginolyticus*.

We measured the flagellar rotation rates of individual free-swimming cells of *V. alginolyticus* as well as their swimming speeds by using laser dark-field microscopy². The figure shows the result obtained at 35 °C. The average value of rotation rate was 1,100 r.p.s. and the highest value was 1,700 r.p.s. The average ratio of swimming speed to rotation rate was 0.113 $\mu\text{m rev}^{-1}$, which means that a cell of *V. alginolyticus* progresses by 7% of the pitch of its flagellar helix (1.58 μm). In addition, we observed a tendency for swimming speed to saturate at high rota-



Swimming speed as a function of flagellar rotation rate obtained for *V. alginolyticus* mutant YM42 (smooth swimming, no lateral flagella). Each point shows data for an individual cell in the presence of 300 mM NaCl at pH 7.0.

tion rate, which we suggest reflects both the torque versus rotation rate characteristics of flagellar motor, and the distribution in lengths of flagella (in preparation).

The sliding velocity between rotor and stator of the flagellar motor rotating at 1,700 r.p.s. is estimated to be about 160 $\mu\text{m s}^{-1}$, because the diameter of the rotor is about 30 nm. Sliding velocities reported for other molecular motors are smaller than this value. For example, myosin moves along actin at a speed of about 10 $\mu\text{m s}^{-1}$ in skeletal muscle and around 100 $\mu\text{m s}^{-1}$ in cytoplasmic streaming of plant cells such as *Chara* and *Nitella*³. So the polar flagella of *V. alginolyticus* should be one of the fastest molecular motors in biological systems.

Our observation of fast flagellar rotation raises some questions about the underlying mechanisms of flagellar motors. The motor of *V. alginolyticus* is driven by Na^+ flow from the outside of the cell to the inside. If we assume that the motor requires 1,000 ions per revolution, as was reported for the H^+ -driven motor of *Streptococcus*⁴, the number of Na^+ ions that move into the cell through the motor per second amounts to about 5% of cytoplasmic Na^+ (concentration of 50 mM; ref. 5). Does fast rotation require such a large influx of ions? Are there any possibilities that the required number of ions per revolution is smaller than 1,000 or reduces at higher rotation rates?

Y. Magariyama, S. Sugiyama

Tsukuba Research Laboratory, Yaskawa Electric Corporation, 5-9-10 Tokodai, Tsukuba 300-26, Japan

K. Muramoto, Y. Maekawa

I. Kawagishi, Y. Imae*

Department of Molecular Biology, Faculty of Science, Nagoya University, Chikusa-Ku, Nagoya 464-01, Japan

S. Kudo

PRESTO, JRDC†

*Deceased.

† C/O Yaskawa Electric Corporation.

1. Lowe, G. *et al. Nature* **324**, 637-640 (1987).
2. Kudo, S. *et al. Nature* **346**, 677-680 (1990).
3. Shimmen, T. *Bot. Mag., Tokyo* **101**, 533-644 (1988).
4. Meister, M., Lowe, G. & Berg, H. C. *Cell* **49**, 643 (1987).
5. Yoshida, S. *et al. J. Biol. Chem.* **265**, 20346 (1990).

The genomes business

Horace Freeland Judson

Correcting the Code: Inventing the Genetic Cure for the Human Body. By Larry Thompson. *Simon and Schuster*: 1994. Pp. 378. \$23.

Dealing with Genes: The Language of Heredity. By Paul Berg and Maxine Singer. *University Science Books*: 1992. Pp. 269. \$34, £17.95. Distributed in the UK by Blackwell Science.

The Human Genome Project: Deciphering the Blueprint of Heredity. Edited by Necia Grant Cooper. *University Science Books*: 1994. Pp. 360. \$38, £27.95. Distributed in the UK by W. H. Freeman.

The Gene Wars: Science, Politics, and the Human Genome. By Robert Cook-Deegan. Norton: 1994. Pp. 416. \$25.

Signs of Life: The Language and Meanings of DNA. By Robert Pollack. *Houghton Mifflin/Viking*: 1994. Pp. 212. \$19.95, £16.

HOWEVER you consider it, the label 'The Human Genome Project' is of course multiply misleading. As we all know, there's no one human genome, and we are deeply interested in precisely those stretches of the genomes that are not silent and where variation is high. Here we uncover clues to disease; here may be buried our talents. Many genomes other than human ones are necessarily part of the project, from HIV through yeast and fruitfly to mouse and chimpanzee. Comparisons among species are even now yielding provocative clues for immediate problems of developmental biology, as for ultimate questions of evolution. And despite the designation 'project' and the assertions made to the sources of money that the thing can be wrapped up not long past the turn of the millennium, in fact and for excellent reasons the research has no foreseeable end. Extending and exploiting it will occupy scientists and soak up funds well past the middle of the next century. Yet the label promotes the project. Try titling it accurately — for example 'the genomes business'. Falls flat; won't sell.

Among those eagerly exploiting the human genome project are publishers. The obligation that the scientists owe the public, to make clear the importance, content and potential consequences of modern genetic research, is matched, fortunately, by considerable public curiosity. The enterprise of genetics has an allure that goes beyond the science itself and beyond any promise of practical benefit. Along with the complementary claims of sociobiology, the human genome project seems to promise to tell us what we are and how we became what we are. A hundred and more books about these matters have been launched at the general reader in the past twenty years. Some marvel, some teach, some warn most direly. Few are distinguished.

The convenient *terminus a quo* is Macfarlane Burnet's *Genes, Dreams and Realities*, which he published in 1971. At the time, he was taken to be attacking molecular biology — whose practitioners

just then, one recalls, were bumptiously confident that the problems of eukaryotes — development and differentiation — were soon to yield, just as prokaryotes had done in the preceding two decades. A few American molecular geneticists in 1971 were beginning to recombine DNA; some, notably Paul Berg, were speculating about gene therapy. Burnet's book was no polemic, although it had a distinct and personal point of view. Revisited today, it seems calmly sceptical about the likeli-

“Despite the designation
'project' and the assertions
made to the sources of
money that the thing can be
wrapped up not long past
the turn of the millennium,
in fact and for excellent
reasons the research has
no foreseeable end”

hood that the new biology will yield very much very soon by way of therapeutics, clear-eyed and pessimistic about the costs and ethical limitations of such therapy even if it were to prove possible. Burnet was celebrated for his biological intuition. Here he showed why, and in a tone that rarely rose above the conversational. The book now seems prescient, and sets a standard of clarity and restraint.

A recent and altogether different sort of book, but one that also sets a standard for its type, is *The Code of Codes*, an ordered set of essays edited by Daniel J. Kevles and Leroy Hood (Harvard University Press, 1992; for a review see *Nature* 358, 27; 1992). The essays, a baker's dozen, range through the history, science and technology of genetic mapping and sequencing to the chief problems of social policy that the project obliges us to consider.

The authors include scientists, among

them Walter Gilbert, James Watson, Charles Caskey, Nancy Wexler and Hood. The fresher, more valuable papers are those that deal with legal and ethical issues: issues of privacy and the limits of genetic testing, as presented by Dorothy Nelkin; of genetics in the courtroom, considered by Eric Lander; of employment and health insurance, treated by Henry Greely; and of the relation of the genome project to eugenics, argued by Kevles. (I am a consultant to the publisher and wrote an historical essay that appears in the book.) Compendia of this sort can be uneven in quality and coverage: *The Code of Codes* does unusually well.

The present crop of books on these subjects varies widely in type and quality. One is so badly done it must be put in the stocks and pelted with clods as a public warning to others. Three promise to last a while.

Larry Thompson's *Correcting the Code* is about gene therapy. All the journalistic clichés are here. Can one tell the story as a race — for a Nobel prize? Against death? Can one find a scientist great of soul and noble of mien? Can one start not at the beginning but with a dramatic high point? Sure enough, Thompson's first chapter, entitled "A Cure for Ashanti", opens with the press conference where W. French Anderson and colleagues announce that they will begin the first authorized attempt to correct a genetic disease by supplying the missing gene, to an otherwise doomed little girl. A page or two shows what sort of book this is. "Restless reporters, pens and microphones poised," awaited the start of the press conference. The scientists "strode in". French Anderson was "the slight, 5 foot, 9 inch, 150-pound chief" of a research unit; he "burned with an uncontained fervour to begin human gene therapy"; now, "a former track star, [he] could see the finish line"; he began speaking, "his silver hair shining in the floodlights".

The treacle flows onwards, but worse is what the treacle smothers. With due respect to French Anderson, he appears to be the willing victim — yet it must be embarrassing to him to be cast as the all-but-exclusive central figure in the development of genetic therapy. Other centres and other scientists have been working on parallel lines, and the fundamental story is of the network of collaboration and competition among them. Beyond that, Thompson oversimplifies the science to the point of incomprehensibility, and repeatedly gets both science and history egregiously wrong. One quick example: describing the discovery of reverse transcriptase, independently by David Baltimore and by Howard Temin and Satoshi Mizutani, he misses the point of Temin's research with Rous sarcoma virus from the early 1960s on — for, in fact, only when Mizutani joined him in the autumn

of 1969 did the work begin that demonstrated the presence of the enzyme. Nor did Temin (or Baltimore) "isolate" the enzyme as Thompson writes. A page later, he asserts that Baltimore, at 35 in 1975, was "the youngest person ever to win the Nobel". No! That was and remains Lawrence Bragg, who had the prize in his pocket in 1915 at the age of 25, while many others had received theirs younger than 35, for example Paul Dirac (at 31), Tsung-Dao Lee (31) and Watson (34).

There has been a need for a competent primer in modern genetics, written for the serious nonspecialist: Paul Berg and Maxine Singer have provided it. They collaborated earlier on a textbook, *Genes and Genomes*. *Dealing with Genes* began as a derivative, but in the writing evolved independently. The result still looks and feels like a textbook, in shape, design and illustration scheme. Their prose is without personality — but that's deliberate, and the book is clear and thorough. It comes in at the right level: it dispenses with the stupefying detail that the student requires, and as a result the bone-structure of the science shows through. The material that is here will age more gracefully than does the full-scale textbook.

The same publisher, University Science Books, has now issued an ambitious, complex and in some ways weird book, which as its title says bears down on the human genome project. This, again, looks like a textbook — a flashy one, large pages printed on stock of various colours and with lots of full-colour artwork and photographs, one even in 3D. It is also a compendium, and of disparate materials by diverse hands — some 50 authors all told. We are told that Necia Grant Cooper, its editor (compiler?) is trained in the physical sciences and works at Los Alamos. A 65-page introduction to genetics is by a

single contributor, but is presented for the most part in a series of 19 full-page 'boxes', in pure US textbook style. This is followed by a 'round table' in which nine prominent scientists, from Baltimore to Wexler and Norton Zinder, wax sententious about "the vision, the science, the implementation" of the genome project, in brief, alternating, heavily edited bursts punctuated by mug shots of the participants. Yet these are followed by four straightforward chapters on the methods, aims and implications of the project. A later chapter, about a specialized new method of sequencing, is five pages long and lists 16 authors. The whole thing is as jumpy and overproduced as a rock-music video, which perhaps defines the generation at which it is aimed.

(University Science Books is worth keeping in mind. Located in Marin County, just north of San Francisco, it is a small house, whose founder, Bruce Armbruster, apparently intends that it should grow into the large, empty niche left in science publishing on the West Coast by the relocation of W. H. Freeman, some years ago, to New York.)

The genome project had all sorts of scientific origins and has been from the beginning the product of politics within science and the interaction of science with politics on the larger scene, in California and Massachusetts, in England and France, but especially in Washington and its surroundings. No one person is in a position to write about all of this firsthand, but Robert Cook-Deegan comes as close as any. He is a physician by training and not a scientist. For crucial years in the mid-1980s, he worked at the Office of Technology Assessment (OTA), a creature of the US Congress, set up to conduct extended studies, at the request of committee chairmen, of anything from, say, coal mining to ozone depletion. There, Cook-Deegan directed projects including one on gene therapy and another, the big one, on the genome project. "Our task at OTA was to report to Congress on what the genome project was, whether or not it was worth funding, and how its underlying bureaucracy should be constructed," he writes. "This vantage point placed me in the midst of three groups: scientists debating the merits of genome projects, administrators working in federal agencies responsible for funding the scientists, and members of Congress who had the constitutional authority to specify how much and in what way federal funds could be used." Cook-Deegan's team was tiny, clever and agile; as the OTA is not part of the federal government they were mostly free of interagency rivalries. Their report, *Mapping Our Genes — Genome Projects: How Big? How Fast?*, issued in 1988, was itself crucial to the founding of the project. And he got to observe most of the players and much of the action. He has

mastered the documents and conducted scores of interviews.

The Gene Wars is a scrupulous and immensely particular narrative of the genesis and opening years of the genome projects. Cook-Deegan is cogent about the balance of scientific, biomedical and commercial motivations, since he is copious in narrating people's actions yet cautious in judging them. Deliciously dry, for example, is the passage about the interactions of Watson and Bernadine Healy, when he was coordinator of genome research for the National Institutes of Health and she the institutes' director manoeuvring to force him out. *The Gene Wars* is often intricate, sometimes stylistically lumpy, but it is an essential history.

To marvel, to teach, to warn — the one recent book that does all three, and with an individual voice and cool wisdom, is Robert Pollack's *Signs of Life*. Pollack is a molecular and cell biologist of deserved reputation, and an extraordinary teacher. Many scientists are thoughtful and widely read outside their specialties: Pollack is of the rare kind whose science itself is permeated and in part shaped by a profound historical, literary and critical sensibility. He is also one of biology's great worriers. It was his telephone call from Cold Spring Harbor to Berg at Stanford in the summer of 1971, about the potential dangers in certain experiments in the then-infant field of recombinant DNA, which set in motion the train of events leading to the moratorium on such research in the summer of 1974 and the Asilomar conference in the spring of 1975. In *Signs of Life*, Pollack's controlling metaphor is that genomes are texts written in DNA, that cells read, edit and select from these texts, and that the grand enterprise of genetics is to read, understand, interpret and perhaps ourselves edit these texts. In his ingenious and supple elucidation of his science, the metaphor works. This is a book that welcomes readers in and, by the end, invites them to consider the limitations of the science, some technical, some prudential, some ethical. Pollack thus rises to the level of Macfarlane Burnet and Peter Medawar: *Signs of Life* is the most distinguished book about science I have seen so far this decade. It is also a sensuous delight to read. □

Horace Freeland Judson is in the Program in History and Philosophy of Science, Stanford University, Stanford, California 94305, USA.

■ Also recently published is *The Book of Man: The Quest to Discover Our Genetic Heritage* by Walter Bodmer and Robin McKie (Little, Brown, £18.99). The authors chart the origins and progress of the Human Genome Project, and provide an optimistic assessment of the potential implications of this research.

New in paperback

The Red Queen: Sex and the Evolution of Human Nature by Matt Ridley.

Penguin, £7.99. "There is a wealth of information here and [the book] is an excellent source for researchers because of the descriptions of studies and its extensive reference section, as well as being interesting to a scientifically literate public. Entertaining... the quality of the writing only rarely falters" (Magnus Enquist and Risa Rosenberg, *Nature* **367**, 697; 1994).

Models for Embryonic Periodicity by Lewis I. Held Jr. Karger, £15, \$24, SFr30, DM36. "Held has brought his incisive intelligence and background in computer programming to bear on a whole menagerie of models. He has scoured the literature with evident enthusiasm. The result is a very useful guide" (Peter A. Lawrence, *Nature* **358**, 720; 1992).

More long arguments

Christopher Wills

Evolution Extended: Biological Debates on the Meaning of Life. Edited by Connie Barlow. MIT Press: 1994. Pp. 333. \$24.95, £22.50.

CONNIE Barlow has followed her successful anthology *From Gaia to Selfish Genes* with another handsomely designed volume of writings by distinguished evolutionists, ranging from Charles Darwin through Julian Huxley to E. O. Wilson. In this compendium she traces their attempts to wrestle with the ineffable in order to find a larger meaning in their work. The ineffable wins, two falls out of three, but in general the evolutionists put up a brave battle.

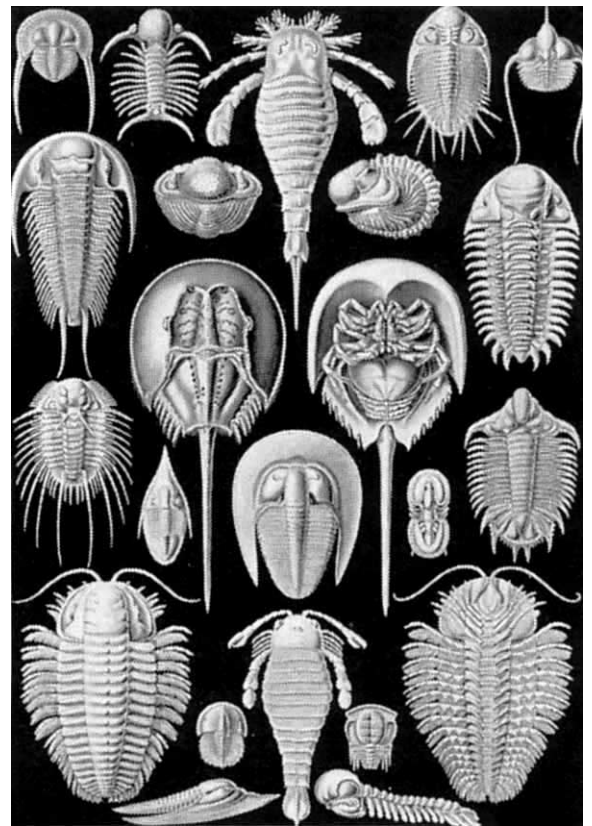
The book will provide the student or interested lay reader with a valuable introduction to many old controversies and a few new ones. Prefacing the various extracts with well-constructed comments, Barlow presents the essential kernels of various arguments about such ineffabilities as evolutionary progress, higher-order and cosmic evolution and the interaction between evolution and religion. Interesting pictures dot the book as well, although she uses rather too many drawings by Gustave Dore of soulful chaps floating about in nightshirts.

The student and lay reader wandering through these pages will find tantalizing fragments of ideas that will jar him or her into further thought. Jacob Bronowski, for

example, is represented by his idea of stratified stability — that evolution cannot go from higher to lower levels of complexity because of a kind of ratchet effect. Mutations that reduce the complexity of the individual might not destroy the individual, but might damage the complex set of associations of which the individual is a part. This is an interesting attempt to explain the ever-increasing ecological complexity that is a hallmark of the history of life.

Sometimes, alas, the ruminations of evolutionists are not helpful in closing the intellectual gaps in our society. Creationist Henry Morris makes an appearance with a gleeful diatribe against Julian Huxley, who had innocently suggested that UNESCO (United Nations Educational, Scientific, and Cultural Organization) should embrace the principles of humanism. Morris accuses Huxley of trying to establish humanism as a worldwide secular religion, hardly what Huxley had in mind. And the less liberal tendencies of evolutionists are on display as well, unfortunately. Peter Medawar, patrician nostrils flaring, inveighs against “a large population of people, often with well developed scholarly and literary tastes, who have been educated far beyond their capacity to undertake analytical thought”. But generally the scientists come off pretty well in comparison to the nonscientists. The most appalling excerpt in the book is taken from Justice Scalia’s dissent from the US Supreme Court decision that invalidated the Louisiana Creationism Act. I have not read enough legal prose to know whether such tendentiousness is the norm—if it is, then Mr Bumble was surely right.

These readings provide a valuable glimpse into the thinking of past decades about these large questions, but there is little included about the greater meaning of the latest startling advances in our understanding of evolution and how it works. This is not surprising. As William Provine points out in a remark included in the book, there seems to be no new generation of deeply religious and thoughtfully philosophical scientists to take the place of Teilhard

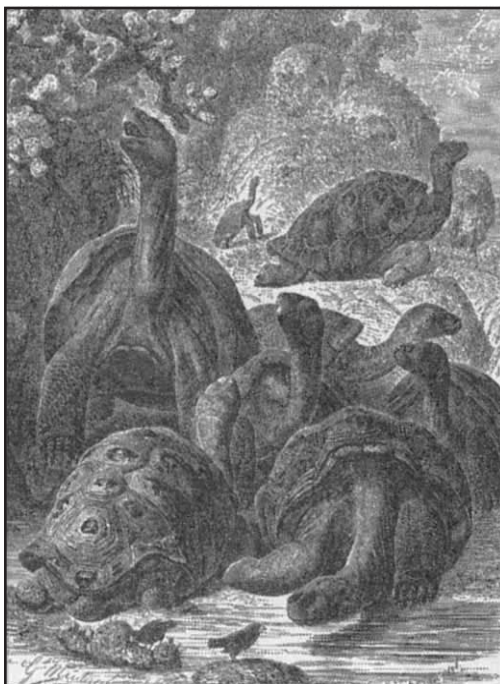


Trilobites, eurypterids and horseshoe crabs, from Ernst Haeckel's *Art Forms in Nature* (1904).

de Chardin, Dobzhansky and Fisher. The increasing secularization of science has been driven by its very success—if one can do an experiment that used only to be in the province of God then where does that leave God? Of course, when these miracle working scientists have passed through their philosophopause, we will find out whether this secularization is reversible.

Things got too fuzzy for me towards the end of the book, which deals with the transformation of science into myth. But many of the themes covered earlier are worth reading in abbreviated form even by busy scientists. For example, there is no doubt that our growing understanding of the Universe is the only thing that keeps our overcrowded society from collapsing. Is there a way to combine this knowledge with an ethical system? Perhaps, although the efforts in this direction by Teilhard, Julian Huxley and others have hardly swept through the world with the force of revelation. By following these scientists' tortuous apologetics, the book shows vividly why building an ethical system based on evolutionary principles is difficult. After all, Leo Szilard's maxim “Do not destroy what you cannot create” is just as difficult to follow in practice as Jesus's maxim “Love thy neighbour as thyself”. □

Christopher Wills is in the Department of Biology, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0116, USA.



Galapagos tortoises and finches, from Alfred E. Brehm's *Merveilles de la nature* (1893).

Name that asteroid

David Hughes

Dictionary of Minor Planet Names, Second Edition. By Lutz D. Schmadel. Springer: 1994. Pp. 741. DM118, \$69, £46.50.

BETWEEN the orbits of Mars and Jupiter lie uncountable numbers of minor planets, or asteroids — to give them their more prosaic modern title. Asteroid number one, Ceres, was discovered on New Year's Day 1801. A thousand asteroids were known by 1921 and the number reached 2,000 by 1946. Over the past decade or so asteroids have been discovered at the rate of about 160 a year. By September last year 5,655 asteroids had been numbered and of these 4,512 had been named.

Numbers are assigned to asteroids when their orbits are known to sufficient accuracy that they can be found at any time in the future. Names are a different matter. The person who discovers an asteroid has ten years to produce a name. There are very few restrictions on poetic inventiveness. The Minor Planet Names Committee of Commission 20 of the International Astronomical Union insists that the name should be pronounceable, preferably expressible as a single word and contain fewer than 16 characters. Names glorifying people or events of principally political or military significance are

considered unsuitable, unless at least a hundred years have elapsed since the person died or the event took place. With these exceptions anything is fair game.

Mythological names predominated in the first three-quarters of the nineteenth century, but today these classical names are almost exclusively reserved for asteroids in a 1:1 resonance with Jupiter. These asteroids commemorate the Trojan war, the Greek besiegers preceding Jupiter and their Trojan opponents following.

At first, all the asteroid names were female and anything that was vaguely masculine was given the feminine suffix of 'a' or 'ia'. Sir Isaac Newton was honoured with asteroid number 662 Newtonia, a German marble became 711 Marmulla, the country of Japan was commemorated with 727 Nipponia and so on. This changed in the early twentieth century and now, among the 4,512 named asteroids, there is a clear preponderance of males to females, in the ratio of 10 to 3.

Some people were suspicious that asteroid nomenclature represented a cemetery for astronomers, but in fact only a quarter of all names come from this profession. Many refer to the relatives of the discoverers. In the case of, for example, Ed Bowell of the Lowell Observatory, Arizona, and Elinor Helin of the Mount Palomar Observatory, California, plenty of relatives are required — they have discovered 341 and 154 asteroids respectively. Among relatives, spouses seem to prevail over both children and parents,

and uncles and aunts are much less represented than the discoverer's friends. Other asteroidal categories are cities, countries, writers, literary figures, composers, plants, animals, painters and acronyms.

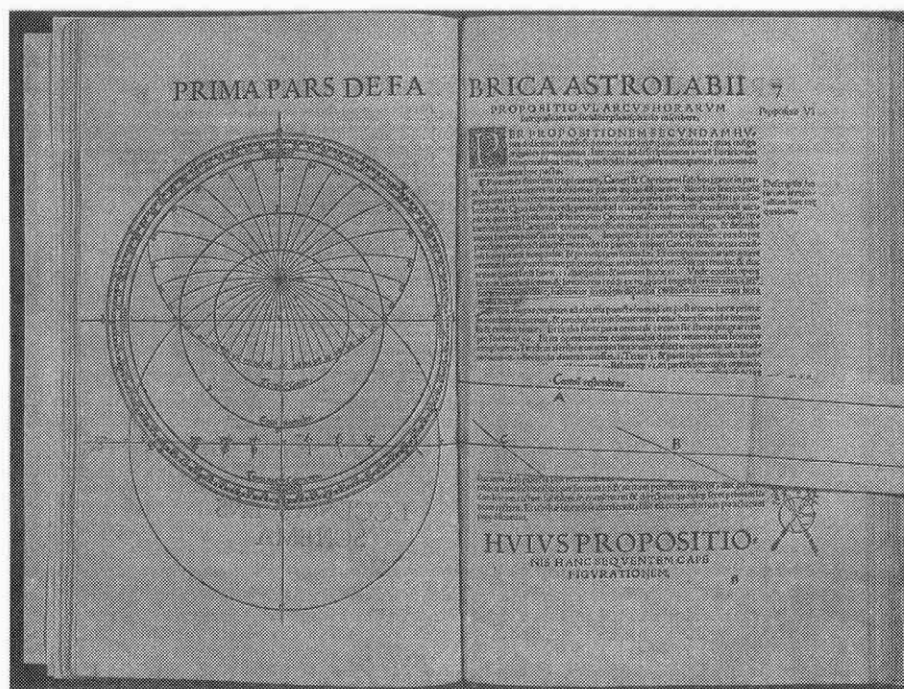
Lutz D. Schmadel has produced an absolutely delightful book. The first edition of this rather specialized compilation was produced in late 1992 and sold out within a few months, to the surprise of both the author and the publisher.

This second edition is a thoroughly revised and enlarged version. The structure is, however, unchanged. The name, preliminary designation, discovery date, discoverer and discovery place are given for all the asteroids, together with a paragraph explaining the meaning of the name and why it was chosen. Since 1947, each name assignment has been accompanied by an explanation, but for many of the 2,000 asteroids named before then, some detailed research had to be done. Even now, Schmadel lists 134 asteroid names with unknown meanings or attributions, most of these asteroids having been discovered by Maximilian Wolf and Karl Reinmuth at Heidelberg and M. Charlois at Nice.

Many happy hours can be spent browsing through this enthralling dictionary. □

David Hughes is in the Department of Physics, University of Sheffield, Hicks Building, Sheffield S3 7RH, UK — and is also asteroid 4206.

The Crawford treasures: a heavenly library in Edinburgh



Johannes Stöfler (1452–1531) achieved fame for calculating 50 years of ephemerides (calender of the daily movement of the planets), but today he is more widely admired for this richly illustrated book on the making and use of an astrolabe. The work, which spawned a host of imitators, is currently on display in "A Heavenly Library: Treasures from the Royal Observatory's Crawford Collection", an exhibition at the National Museums of Scotland, Edinburgh (until 31 December). The collection was put together over a century ago by James

Ludovic Lindsay, 26th Earl of Crawford, an amateur astronomer and a bibliophile, and stands today as one of the finest libraries of rare astronomical books. An exhibition catalogue is available from the observatory at £12.50 (pbk).

Distance to the Virgo cluster galaxy M100 from Hubble Space Telescope observations of Cepheids

Wendy L. Freedman^{*}, Barry F. Madore[†], Jeremy R. Mould[‡], Robert Hill^{*}, Laura Ferrarese^{§¶}, Robert C. Kennicutt Jr^{||}, Abhijit Saha[¶], Peter B. Stetson^{##}, John A. Graham^{**}, Holland Ford^{§¶}, John G. Hoessel^{††}, John Huchra^{‡‡}, Shaun M. Hughes^{§§} & Garth D. Illingworth^{|||}

^{*} Carnegie Observatories, 813 Santa Barbara Street, Pasadena, California 91101, USA

[†] NASA/IPAC Extragalactic Database, Infrared Processing and Analysis Center, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91125, USA

[‡] Mount Stromlo and Siding Spring Observatories, Private Bag, Weston Creek Post Office ACT 2611, Sydney, Australia

[§] Department of Physics and Astronomy, Bloomberg 501, Johns Hopkins University, 3400 North Charles Street, Baltimore, Maryland 21218, USA

[¶] Steward Observatory, University of Arizona, Tucson, Arizona 85721, USA

[¶] Space Telescope Science Institute, Homewood Campus, Baltimore, Maryland 21218, USA

^{##} Dominion Astrophysical Observatory, 5071 West Saanich Road, Victoria, British Columbia, V8X 4M6, Canada

^{**} Department of Terrestrial Magnetism, Carnegie Institution of Washington, 5241 Broad Branch Road North West, Washington DC 20015, USA

^{††} Department of Astronomy, University of Wisconsin, 475 North Charter Street, Madison, Wisconsin 53706, USA

^{‡‡} Harvard-Smithsonian, Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts 02138, USA

^{§§} Royal Greenwich Observatory, Madingley Road, Cambridge CB3 0HA, UK

^{|||} Lick Observatory, University of California, Santa Cruz, California 95064, USA

Accurate distances to galaxies are critical for determining the present expansion rate of the Universe or Hubble constant (H_0). An important step in resolving the current uncertainty in H_0 is the measurement of the distance to the Virgo cluster of galaxies. New observations using the Hubble Space Telescope yield a distance of 17.1 ± 1.8 Mpc to the Virgo cluster galaxy M100. This distance leads to a value of $H_0 = 80 \pm 17$ km s⁻¹ Mpc⁻¹. A comparable value of H_0 is also derived from the Coma cluster using independent estimates of its distance ratio relative to the Virgo cluster.

WITHIN the framework of general relativity, the evolution of the Universe can be specified by the Friedmann equation which relates the expansion rate H , to the mean mass density ρ , the curvature k , and a possible additional term, called the cosmological constant Λ (identified with the gravitational effects of the vacuum energy density). In a uniform and isotropic Universe the relative expansion velocity v is proportional to the relative distance r such that $v = H \times r$. Thus a determination of the present-day value of the Hubble constant H_0 determines both the expansion timescale and the size scale of the Universe. The Hubble constant also provides constraints on the density of baryons produced in the Big Bang, the amount of dark matter, and how structure formed in the early Universe^{1,2}.

Despite 65 years of study, the value of the Hubble constant has remained in dispute. Although the measurement of relative velocities of galaxies is straightforward, the measurement of accurate distances has always been more difficult. For distances out to ~ 5 Mpc, there is now general agreement^{3,4} to a level of better than $\sim \pm 10\%$. However, for distances as great as the Virgo cluster there is a significant discrepancy^{5,6}, with quoted values ranging from about 15 to 24 Mpc. Not only is there a range of published distances, but there is a tendency for the values to cluster at the extremes, giving rise to the so-called 'short' and 'long' distance scales.

The most accurate means of measuring the distances to galaxies has proved to be the application of a relationship between the period and the luminosity for a class of supergiant variable stars known as classical Cepheids. In fact, Cepheids were used by Edwin Hubble to demonstrate the extragalactic nature of the spiral nebulae^{7,8}. They are relatively young, massive stars with

luminosities $\sim 1,000$ to $100,000$ times brighter than the Sun ($-2 < M_V < -7$ mag). They are well calibrated; they are easy to identify in external galaxies because of their variability and brightness; and they have measured dispersions in the V - and I -band period-luminosity (P - L) relationships amounting to only ± 0.25 and ± 0.20 mag, respectively³. In addition, the underlying physical basis for the Cepheid P - L relation is well understood.

Unfortunately, Cepheids are not luminous enough to be observed out to distances where galaxies are participating in the free expansion of the Universe. The motions of individual galaxies can be perturbed by interactions with nearby neighbouring galaxies; in addition, galaxies can participate in large-scale flows^{9,10}. Hence, to measure the pure Hubble flow, other (secondary) distance techniques must be used to extend the extragalactic distance scale beyond the observable range of the Cepheids, to recession velocities of a few thousand km s⁻¹ where peculiar velocities (which can amount to several hundred km s⁻¹) are small in comparison to the expansion velocity.

The Virgo cluster of galaxies is close enough to be studied in detail; and yet it is far enough away that it is of cosmological interest. It is rich in both spiral and elliptical galaxies and therefore it has played a critical role in the extragalactic distance scale and determination of the Hubble constant. Thus obtaining a Cepheid distance to the Virgo cluster represents a crucial step in resolving the current uncertainty in H_0 and resolving the dichotomy in distance estimates. A variety of other techniques for measuring distances have been applied to the Virgo cluster; hence, an accurate direct measure of the distance to this cluster can be used to calibrate (or set the zero point for) other secondary distance indicators.

The key project

The goal of the Hubble Space Telescope Key Project on the extragalactic distance scale is to provide a measure of the Hubble constant accurate to 10%. This aim is non-trivial given that the history of previous attempts to measure the extragalactic distance scale is replete with examples where large systematic errors were eventually revealed. Hence, the determination of accurate distances requires careful attention to eliminating potential sources of systematic error.

The strategy adopted by the Key Project team on the extragalactic distance scale is threefold. The first goal is to discover Cepheids, and thereby measure accurate primary distances to spiral galaxies located in the field and in small groups that are suitable for the calibration of several independent secondary methods. (These secondary methods include: the Tully–Fisher relationship^{11,12}, the surface-brightness fluctuation method¹³, the planetary nebula luminosity function¹⁴, the expanding photosphere method applied to type II supernovae^{15,16}, and the measurement of the luminosities of type Ia supernovae^{6,17–19}.) The second objective is to provide a check on potential systematic errors in the Cepheid distance scale through independent distance estimates to the nearby galaxies, M31, M33 and the Large Magellanic Cloud (LMC) and, in addition, to undertake an empirical test of the sensitivity of the zero point of the Cepheid P – L relationship to heavy-element abundances. The third and most challenging observational goal is to make Cepheid measurements of distances to three spiral galaxies in the Virgo cluster and two members of the Fornax cluster.

The distance to M100

Our first observations aimed at finding Cepheids in a Virgo cluster galaxy were made in a two-month period beginning in April

1994 with a sequence of 12 1-hour V -band exposures of a field ~ 2 arcmin east of the nucleus of M100. The observing strategy was designed to provide well-sampled light curves for Cepheids having periods ranging from 10 to 60 days. In addition, I -band exposures were taken back-to-back with 4 of the V observations. We present here sample light curves for the Cepheids discovered, their V -band and I -band P – L relationships, and a preliminary estimate of the distance to M100. Details of the data reduction and analysis, tabulation of the photometry, and identification of the variables will be presented elsewhere (L.F. *et al.*, manuscript in preparation; R.H. *et al.*, manuscript in preparation). Three independent calibrations of the photometric zero points were made; the final resulting uncertainties in the V and I zero points amount to ± 0.05 and ± 0.04 mag, respectively. A description of the sampling strategy used for the optimal discovery of variables, and of the method used for identification of variables, determination of mean magnitudes, reddening, and distance have been published elsewhere²⁰.

V -band light curves for twelve M100 Cepheids are illustrated in Fig. 1. The total sample of Cepheids has a range of periods from 20 to 65 days, and mean V magnitudes ranging from 25.0 to 26.5 mag. It is evident from the small scatter in the light curves that the quality of the photometry being obtained by the Hubble Space Telescope (HST) is excellent. Internal estimates of the random errors for single data points are found to be ± 0.14 mag at $V \approx 25$ mag and ± 0.17 mag at $V \approx 26$ mag. The errors in the mean magnitudes for the 12 epochs amount to ± 0.04 and ± 0.05 mag at these same magnitude levels.

In Fig. 2, we present the time-averaged, intensity-weighted V -band and I -band P – L relations for 20 Cepheids in M100 (white circles). We compare the brightness of the M100 Cepheids with those in the LMC (black circles) to obtain a distance ratio

TABLE 1 Error budget in the M100 distance scale

Source of uncertainty	Type of uncertainty	Error (mag.)	Cumulative error
Distance to M100			
[A] LMC Distance	Independent estimates	± 0.10	
[B] WFPC2 V -band zero point	Ground-based/on-orbit calibration	± 0.05	
[C] WFPC2 I -band zero point	Ground-based/on-orbit calibration	± 0.04	
[D] Extinction-corrected Cepheid modulus	[B] and [C] are uncorrelated	± 0.17	
[E] Charge transfer efficiency	4% across entire chip	± 0.01	
[F] Cepheid V -band modulus	(V -band P – L dispersion)/ $\sqrt{N}$	± 0.06	
[G] Cepheid I -band modulus	(I -band P – L dispersion)/ $\sqrt{N}$	± 0.07	
[H] Cepheid true modulus	[F] and [G] are correlated	± 0.03	
[I] = [A] + [D] + [E] + [H]	Combined in quadrature		± 0.20
Distance to Virgo			
[J] Velocity of Virgo cluster	$\pm 80 \text{ km s}^{-1}$	± 0.11	
[K] Back-to-front geometry of Virgo	$\pm 3 \text{ Mpc}$	± 0.35	
[L] = [I] + [J] + [K]	Combined in quadrature $H_0 = 82 \pm 17 \text{ km s}^{-1} \text{ Mpc}^{-1}$		± 0.42
Distance to Coma			
[M] Velocity of Coma cluster	$\pm 100 \text{ km s}^{-1}$	± 0.02	
[N] Virgo/Coma Relative distance	Secondary indicators	± 0.10	
[O] = [I] + [K] + [M] + [N]	Combined in quadrature $H_0 = 77 \pm 16 \text{ km s}^{-1} \text{ Mpc}^{-1}$		± 0.42

A quantitative overview of the propagation of errors associated with our new data set and its application to the extragalactic distance scale. Errors in the photometric zero points and statistical uncertainties associated with the mean apparent distance moduli both propagate into the determination of a true distance modulus. In the case of independent errors in the photometric calibration the terms are assumed to be fully uncorrelated giving

$$\varepsilon_{\mu_0}^U = R_V(\varepsilon_V^2 + \varepsilon_I^2)^{1/2}$$

where $R_V = A_V/E(V-I) = 2.6$ is the ratio of total-to-selective absorption. For the individual wavelength-dependent moduli the uncertainties are correlated in sign and magnitude because in this case the Cepheid samples defining the means are identical, thereby giving

$$\varepsilon_{\mu_0}^C = R_V(\varepsilon_V - \varepsilon_I)$$

As can readily be seen, uncertainties in the photometric zero points dominate the error budget for the calculated distance to M100; with additional ground-based observing this error can be significantly reduced.

(M100/LMC). In logarithmic units of magnitudes, the so-called 'distance modulus' (μ) is given by $5 \times \log_{10}$ (distance in parsecs) $- 5.0$. A correction for the dimming effect due to interstellar dust is obtained by fitting the apparent V and I moduli to a standard interstellar extinction law. The apparent I -band modulus to M100 is measured to be $\mu_I = 31.25$ mag; the corre-

sponding V P - L relation for the same sample of Cepheids yields $\mu_V = 31.31$ mag, implying a mean extinction of $A_V = 0.15 \pm 0.17$ for the M100 Cepheid sample. The true (reddening-corrected) distance modulus to M100 is determined to be 31.16 ± 0.20 mag, corresponding to a distance of 17.1 ± 1.8 Mpc. A detailed listing of the errors for this determination is given in Table 1. The

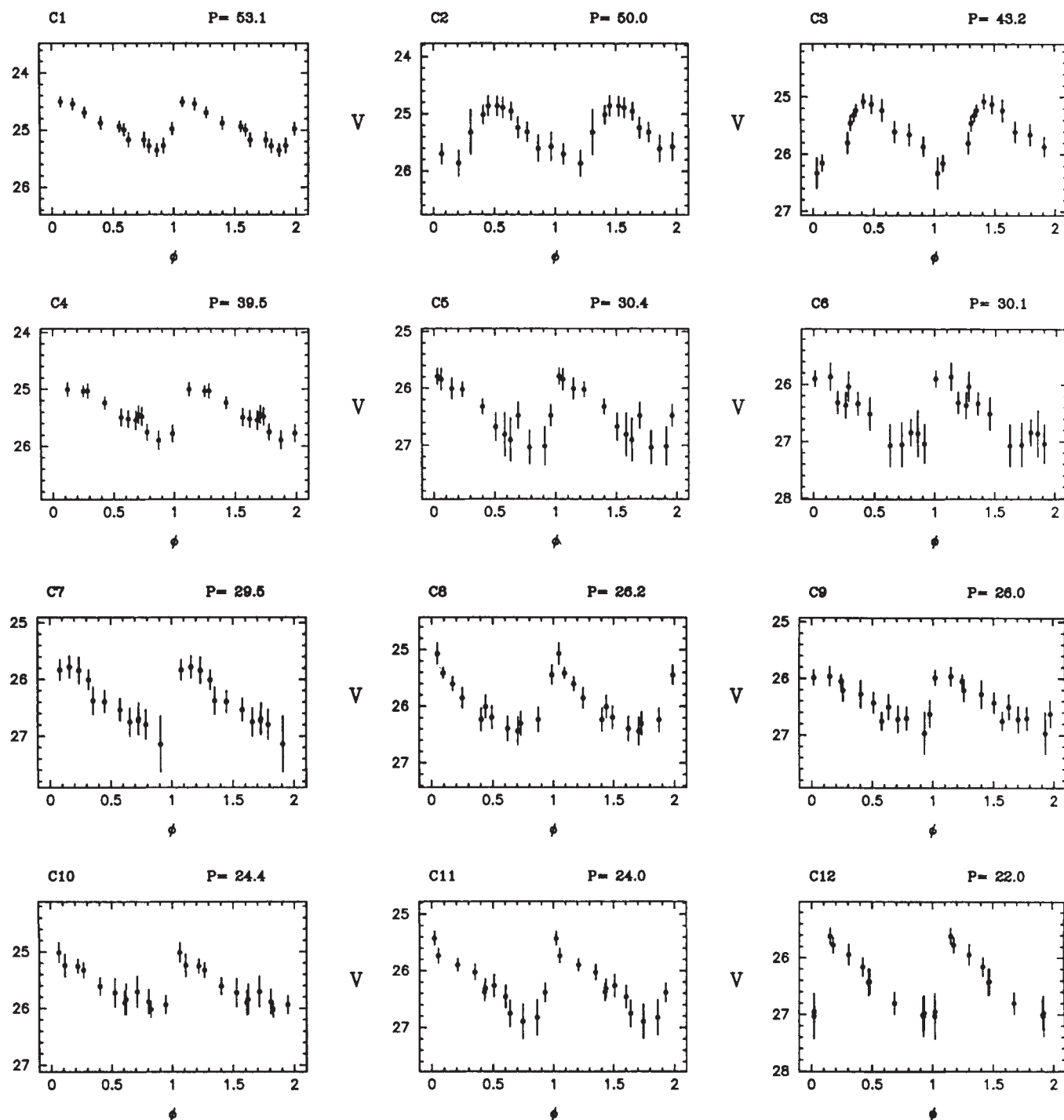


FIG. 1 Sample light curves for twelve Cepheids in M100. The periods of the stars illustrated (~ 20 to 50 days) are typical of the total sample discovered in M100. In addition to their periods, Cepheids are characterized by their distinctive light curves showing a rapid rise to maximum light, followed by a slower linear decline to minimum light. Cepheids pulsate because their atmospheres are not in hydrostatic equilibrium. The instability occurs as a result of a change in the opacity of a helium

ionization zone with temperature. As singly ionized helium becomes doubly ionized, the opacity increases with increasing temperature. This changing opacity acts as a valve so that when the atmosphere is compressed and at higher temperature, it becomes more opaque to radiation, and the thermal energy increases. Subsequently, it expands and cools, becoming more transparent. As the pressure on the atmosphere decreases it then collapses and the cycle repeats.

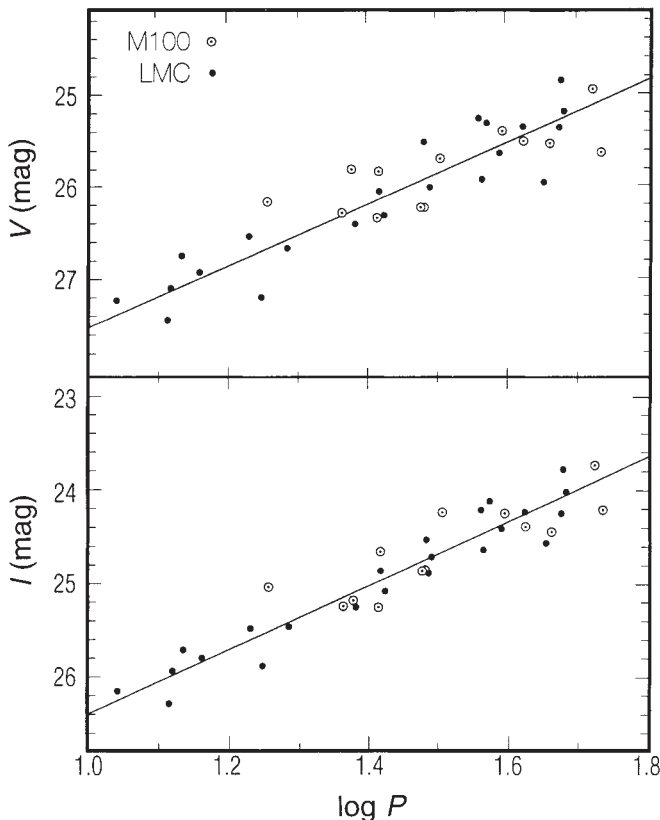


FIG. 2 Composite period-luminosity relations (V-band upper panel; I-band lower panel) for Cepheids in M100 (white circles) and Cepheids in the LMC (black circles, shifted to the distance of M100). Apparent V and I distance moduli for M100 are obtained by minimizing the residuals in the combined $P-L$ relations for the two galaxies, and determining the relative offset with respect to the calibrating LMC sample. Consistent with previous studies^{3,20}, only high signal-to-noise variables (those having an average of their absolute deviations from the mean exceeding 1.5 times the mean error) are plotted, whereas only stars with $\log P < 1.8$ are included in the fit. The difference in the V and I apparent moduli for M100 is assumed to be due to interstellar dust present both in M100 and our own galaxy. Correcting for this effect yields a reddening-corrected (true) distance to M100 of 17.1 ± 1.8 Mpc.

random errors in the apparent moduli amount to $<4\%$ in distance. The uncertainty in the true distance is dominated by systematics due primarily to the adopted LMC distance and correction for reddening, giving a total uncertainty of $\pm 10\%$.

This first Cepheid distance to M100 falls toward the low end of the published range of distance values to the Virgo cluster. How does this distance compare with previous distance estimates to M100 itself? We discuss four recent measurements of the distance to M100. The first measurement is based on a comparison of the observed diameters of the Scl galaxies. Assuming that M100 has the same diameter as M101, Sandage²¹ concluded that M100 is located at a distance of 27.7 Mpc. Ten years earlier, de Vaucouleurs²² discussed the relative positions of 37 spirals in the Virgo cluster based on a variety of secondary distance indicators available at that time. He found a distance to M100 of 11.8 Mpc, which places M100 on the near side of the Virgo cluster which he estimated to be at a distance of 15.0 Mpc. A third estimate of the distance to M100 is obtained from the expanding photosphere method^{15,16} applied to the type II supernova SN1979c. The most recent application of this method yields a distance of 15 ± 4 Mpc. As discussed further (J.R.M. *et al.*, manuscript in preparation), the distances to the four galaxies now available using both Cepheids and type II supernovae all show good agreement. Fourth, M100 is contained in the Virgo galaxy sample

studied by Pierce and Tully²³. On the basis of the Tully-Fisher (TF) relationship, these authors conclude that M100 lies at a distance of 18.4 ± 2.2 Mpc, which they claim is within 1σ of their mean distance to the Virgo cluster (15.6 ± 1.5 Mpc). A somewhat closer distance of 14.5 ± 2.7 Mpc is given in Table 3 by Pierce²⁴. The M100 Tully-Fisher and type II supernova estimates are in good agreement with the distance derived by Cepheids.

Where does M100 lie with respect to the centre of the Virgo cluster? An answer to this question is needed to determine H_0 from the present data. M100 is projected on the sky 3.9° from the centre of the Virgo cluster as defined by the giant elliptical galaxy M87. The angular extent of the core is about 6° radius, corresponding to 1.8 Mpc at a distance of 17.1 Mpc. The full front-to-back depth of the elliptical galaxy component of the cluster has been estimated to be 8 Mpc (ref. 25). This range would imply an r.m.s. variation in depth of ± 2 Mpc for the elliptical galaxies, if the radial distribution is gaussian. However, the distribution in depth of the spiral galaxies in the cluster appears to be larger, with a full range of ~ 13 Mpc (ref. 23), and this distribution is more complex than a simple gaussian model. With these uncertainties in mind we conservatively assign an error of $\pm 20\%$ in the mean distance of the Virgo cluster. The confidence limits that can be placed on this Virgo distance will be discussed in more detail (J.R.M. *et al.*, manuscript in preparation).

By contrast, is there any reason to believe that M100 lies in the foreground? Sandage and Bedke²⁶ present a list of 75 galaxies in the direction of the Virgo cluster that resolve most easily into individual stars. On the basis of their resolution, 45 of these galaxies are classified as 'excellent', 'good' or 'good-to-fair', whereas M100 itself is classified as 'fair'. It then seems unlikely that M100 is significantly closer than most of the Virgo spiral galaxies discussed by Sandage because resolution is distance dependent. Nevertheless, the line-of-sight position of M100 with respect to the core of the Virgo cluster is the major source of uncertainty in what follows.

There is a second source of uncertainty in estimating H_0 from the Virgo cluster even if the distance were accurately known. The cluster is sufficiently far that its Hubble recession velocity is larger than its peculiar velocity. However, there are many difficulties in determining the Hubble velocity: our own Local Group is infalling toward the Virgo cluster, and the presence of foreground plus background galaxies complicates the determination of the Virgo cluster velocity itself. The range of corrections for the Local Group centroid, the adopted infall velocity and the random velocity can lead to Virgo recession velocities²⁷ from 1,200 to 1,600 km s^{-1} . Adopting the recession velocity of $1,404 \pm 80 \text{ km s}^{-1}$ preferred by Huchra²⁷ and a Virgo cluster distance of 17.1 Mpc yields a value of $H_0 = 82 \pm 17 \text{ km s}^{-1} \text{ Mpc}^{-1}$. The cumulative errors entering this estimate are listed in Table 1. (However, we note that adopting the velocity^{6,31} of $1,179 \pm 17 \text{ km s}^{-1}$ results in $H_0 = 69 \pm 14 \text{ km s}^{-1} \text{ Mpc}^{-1}$ for the same distance.)

A means of avoiding the uncertainty in the Virgo cluster recession velocity in determining H_0 is to make use of the well measured distance ratio between the Virgo and Coma clusters²⁸. The Coma cluster is located about 6 times more distant than the Virgo cluster. Its peculiar velocity is $\lesssim 80 \text{ km s}^{-1}$ (ref. 29) and it has a recession velocity^{9,30,31} of $7,200 \pm 100 \text{ km s}^{-1}$. Hence its peculiar velocity contributes a very small fractional uncertainty to the cosmological velocity. The Virgo-Coma cluster distance ratio has been measured using a variety of techniques (for example, the Tully-Fisher relationship, the $D_n-\Sigma$ relation, and type Ia supernovae). There is excellent agreement among a number of different authors^{4,32} resulting in a mean relative Coma-Virgo distance modulus of $3.71 \pm 0.10 \text{ mag}$. Adopting a distance of 17.1 Mpc for the Virgo cluster yields a distance to the Coma cluster of 94 Mpc. The corresponding value of the Hubble constant is $H_0 = 7,200/94 = 77 \pm 16 \text{ km s}^{-1} \text{ Mpc}^{-1}$. This value agrees well with values determined for the Virgo cluster velocities

given above. Thus a value of the Hubble constant $\sim 80 \pm 17 \text{ km s}^{-1} \text{ Mpc}^{-1}$ is indicated extending out to a distance of $\sim 100 \text{ Mpc}$.

Systematic errors

Could there be a systematic error in the local calibrators, possibly associated with the zero point of the Cepheid distance scale? This possibility seems very unlikely given that the distances to nearby galaxies have been measured using several different methods completely independent of the Cepheids; for example, using RR Lyrae variables^{33,34}, using the photoionized disk around SN1987A³⁵, the expanding photosphere method applied to SN1987A¹⁵, and using the luminosity of the tip of the red giant branch³⁶. Although as recently as a decade ago the distances to the local calibrators were in disagreement at a level of a factor of two, newer data obtained from CCDs and near-infrared arrays have led to a convergence of the distances to local galaxies with differences amounting to less than 0.3 mag, or 15% in distance³⁷.

Are we measuring the true global value of H_0 ? Given the observational evidence for large-scale velocity flows^{9,38}, and theoretical models for structure formation (such as cold dark matter), it is possible that measurement of H_0 locally does not yield a value representative of that on a larger global scale^{39,40}. However, there are several direct lines of evidence that do not support this hypothesis. At the distance of the Coma cluster such fluctuations are predicted to be less than 5–10%, based both on models and considerations of the observed local bulk motion with respect to the microwave background³⁹. At least three distance indicators presently extend well beyond the distance of the Coma cluster and yield internally consistent values of H_0 both locally and out to redshifts $> 14,000 \text{ km s}^{-1}$. These are the type II supernovae^{15,16}, type Ia supernovae⁶, and

brightest cluster members^{41,42}. Moreover, the Hubble diagram for the brightest cluster members of Sandage and Hardy⁴² is well-defined to velocities $> 40,000 \text{ km s}^{-1}$ and appears to remain linear out to $\sim 100,000 \text{ km s}^{-1}$. The zero-point calibration of these various methods is still a subject of debate (and the goal of the Key Project) but, for example, it is concluded⁴¹ that the variation of H_0 over the redshift range $0.01 < z < 0.05$ is constrained to be within $\pm 7\%$ r.m.s.

Implications

The HST measurement of the Cepheid distance to M100 enables us to place constraints on the range of plausible values of H_0 and the expansion age. The current limits on H_0 are illustrated in Fig. 3. Also shown in Fig. 3 are the expansion ages for various values of (H_0, Ω_0) corresponding to $14 \pm 2 \text{ Gyr}$. These limits are broadly consistent with estimates of the ages of globular clusters from stellar evolution theory^{43–45}, estimates of the age of the Galactic disk based on white dwarf cooling times^{46,47}, and radioactive dating of elements^{48,49}. All of these age estimates generally span a full range between 10 to 18 Gyr, consistent with the 1σ error limits quoted above. A value of $H_0 = 80 \pm 17 \text{ km s}^{-1} \text{ Mpc}^{-1}$ is consistent with a low-density ($0.1 < \Omega < 0.3$) Universe and $T_0 = 12 \text{ Gyr}$.

The Einstein–de Sitter cosmological model (often referred to as the standard model) has a mean mass density equal to the critical density ($\Omega = 1$) and a cosmological constant, $\Lambda = 0$. In the standard model the expansion age is given by $T_0 = 2/3 H_0^{-1}$. (For models with $\Lambda > 0$ the Universe undergoes an accelerating expansion resulting in ages that are older than $\Lambda = 0$ models⁵⁰.) For $H_0 = 80 \text{ km s}^{-1} \text{ Mpc}^{-1}$ the standard model gives an expansion age of 8 Gyr, well below the other age estimates cited above. This ‘age conflict’ suggests that either the standard cosmological model needs to be revised, or present theories (or observations) bearing on stellar and galactic evolution may need to be re-examined.

Concerns and future plans

The weakest point in the present analysis is of course that we have measured a distance to only one galaxy. However, there has recently been reported⁵¹ the discovery of three Cepheid candidates in the Virgo cluster spiral galaxy NGC4571 based on observations taken at the Canada–France–Hawaii Telescope. It is concluded from these data that NGC4571 lies at a distance of $14.9 \pm 1.2 \text{ Mpc}$, yielding a Hubble constant $H_0 = 87 \pm 7 \text{ km s}^{-1} \text{ Mpc}^{-1}$. Clearly, however, the distances to more galaxies are required to define the mean Virgo cluster distance. We do not wish to mislead the reader into believing that the problem of determining H_0 has been solved. It has not. Our analysis has shown that the remaining uncertainty in the value of the Hubble constant is dominated by systematic errors, which are difficult to quantify with a sample of only one galaxy. An accuracy of 10% or better in the value of H_0 will only be reached when we have measured Cepheid distances to a larger sample of galaxies so that the magnitude of these systematic errors can be assessed directly. However, the independent estimate of the distance to NGC4571, the consistency of the type II supernovae and Tully–Fisher distance scales with the direct Cepheid measurement to M100, and the agreement of the distance to the Virgo cluster based on elliptical galaxies^{14,25}, lead us to conclude that the evidence at this time favours a value of $H_0 \approx 80 \pm 17 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

The two remaining years of the Key Project will be critical. We have been awarded time and are currently obtaining data to measure Cepheid distances to a total of ~ 20 calibrating spirals both in the field and in small groups (for example, Leo I and Coma I), and including two additional galaxies in each of the Virgo and Fornax clusters. A direct comparison of the Tully–Fisher, surface brightness fluctuation, planetary nebula luminosity function, and type I and II supernovae (and other) distance scales can then be made. In parallel, independent horizontal-

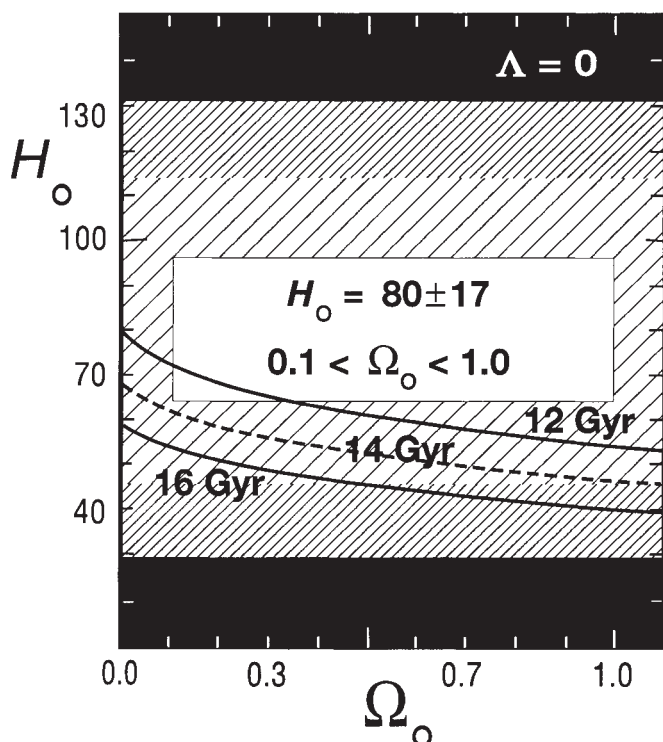


FIG. 3 Current limits on H_0 (in $\text{km s}^{-1} \text{ Mpc}^{-1}$) and the age and density of the Universe. Expansion ages of $14 \pm 2 \text{ Gyr}$, corresponding to globular cluster ages, are illustrated as a function of H_0 and the density parameter Ω for models with the cosmological constant $\Lambda = 0$. The 2σ , 3σ and 4σ confidence intervals on the Hubble constant presented here are shown by the lightly hatched, densely hatched and black horizontal areas, respectively. Broad limits on Ω are illustrated at 0.1 and 1.0.

branch distances to M33 and M31 will be measured, and main-sequence fitting undertaken for LMC clusters, allowing further checks on the zero point of the Cepheid $P-L$ relation. These data will provide a direct means of assessing the systematic errors in the current extragalactic distance scale at each stage of its

application, and allow a determination of the Hubble constant to unprecedented accuracy. The results of this paper suggest that a Hubble constant, accurate to 10%, and measured through secondary techniques out to scales of hundreds of Mpc, is now a realizable goal. □

Received 20 September; accepted 26 September 1994.

1. Fukugita, M., Hogan, C. & Peebles, P. J. E. *Nature* **366**, 309–312 (1993).
2. Kolb, E. W. & Turner, M. S. *The Early Universe* (Addison-Wesley, Redwood City, 1990).
3. Madore, B. F. & Freedman, W. L. *Publ. astron. Soc. Pacif.* **103**, 933–957 (1991).
4. van den Bergh, S. *Publ. astron. Soc. Pacif.* **104**, 861–883 (1992).
5. Jacoby, G. H. et al. *Publ. astron. Soc. Pacif.* **104**, 599–662 (1992).
6. Sandage, A. R. & Tammann, G. A. *Astrophys. J.* **415**, 1–9 (1993).
7. Hubble, E. P. *Astrophys. J.* **62**, 409–443 (1925).
8. Hubble, E. P. *Astrophys. J.* **63**, 236–274 (1926).
9. Aaronson, M. et al. *Astrophys. J.* **302**, 536–563 (1986).
10. Dressler, A. et al. *Astrophys. J.* **313**, L37–L42 (1987).
11. Tully, R. B. & Fisher, J. R. *Astr. Astrophys.* **54**, 661–673 (1977).
12. Pierce, M. J. & Tully, R. B. *Astrophys. J.* **387**, 47–55 (1992).
13. Tonry, J. & Schneider, D. P. *Astr. J.* **96**, 807–815 (1988).
14. Jacoby, J., Ciardullo, R. & Ford, H. C. *Astrophys. J.* **356**, 332–349 (1990).
15. Schmidt, B. P., Kirshner, R. P. & Eastman, R. G. *Astrophys. J.* **395**, 366–386 (1992).
16. Schmidt, B. P., Kirshner, R. P. & Eastman, R. G. *Astrophys. J.* **432**, 42–48 (1994).
17. Branch, D. & Miller, D. L. *Astrophys. J.* **405**, L5–L8 (1993).
18. Phillips, M. M. *Astrophys. J.* **413**, L105–L108 (1993).
19. Saha, A. et al. *Astrophys. J.* **425**, 14–34 (1994).
20. Freedman, W. L. et al. *Astrophys. J.* **427**, 628–655 (1994).
21. Sandage, A. R. *Astrophys. J.* **402**, 3–14 (1993).
22. de Vaucouleurs, G. *Astrophys. J.* **253**, 520–525 (1982).
23. Pierce, M. J. & Tully, R. B. *Astrophys. J.* **330**, 579–595 (1988).
24. Pierce, M. J. *Astrophys. J.* **430**, 53–62 (1994).
25. Tonry, J. L., Ajhar, E. A. & Luppino, G. A. *Astr. J.* **100**, 1416–1423 (1990).
26. Sandage, A. R. & Bedke, J. *Atlas of Galaxies Useful For Measuring the Cosmological Distance Scale*, SP-496 (NASA, 1988).
27. Huchra, J. in *Extragalactic Distance Scale* (eds van den Bergh, S. & Pritchet, C. J.) *Publ. astron. Soc. Pacif. Conf. Series* **4**, 257–280 (1988).
28. Tammann, G. A. & Sandage, A. R. *Astrophys. J.* **294**, 81–95 (1985).

29. Han, M. & Mould, J. R. *Astrophys. J.* **396**, 453–459 (1992).
30. Fukugita, M. et al. *Astrophys. J.* **376**, 8–22 (1991).
31. Jergen, H. & Tammann, G. A. *Astr. Astrophys.* **276**, 1–8 (1993).
32. de Vaucouleurs, G. *Astrophys. J.* **415**, 10–32 (1993).
33. Pritchet, C. J. & van den Bergh, S. *Astrophys. J.* **316**, 517–529 (1987).
34. Saha, A., Freedman, W. L., Hoessel, J. G. & Mossman, A. E. *Astr. J.* **104**, 1072–1085 (1992).
35. Panagia, N. et al. *Astrophys. J.* **380**, L23–L26 (1991).
36. Lee, M. G., Freedman, W. L. & Madore, B. F. *Astrophys. J.* **417**, 553–559 (1993).
37. Freedman, W. L. & Madore, B. F. in *New Perspectives on Stellar Pulsation and Pulsating Stars* (eds Nemec, J. & Matthews, J.) (Cambridge Univ. Press, 1993).
38. Lynden-Bell, D., Lahav, O. & Burstein, D. *Mon. Not. R. astr. Soc.* **241**, 325–345 (1989).
39. Turner, E., Cen, R. & Ostriker, J. P. *Astr. J.* **103**, 1427–1437 (1992).
40. Bartlett, J. G. et al. *Science* (submitted).
41. Lauer, T. & Postman, M. *Astrophys. J.* **400**, L47–L50 (1992).
42. Sandage, A. R. & Hardy, E. *Astrophys. J.* **183**, 743–757 (1973).
43. Renzini, A. in *Observational Tests of Inflation* (eds Shanks, T. et al.), (Kluwer, Boston, 1991).
44. Sandage, A. R. *Astr. J.* **106**, 719–725 (1993).
45. Shi, X., Schramm, D. N., Dearborn, D. S. P. & Truran, J. W. *Science* (submitted).
46. Winget, D. E. et al. *Astrophys. J.* **315**, L77–L81 (1987).
47. Pitts, E. & Taylor, R. J. *Mon. Not. R. astr. Soc.* **255**, 557–560 (1992).
48. Fowler, W. in *14th Texas Symp. on Relativistic Astrophysics* (ed Fenyöes, E. J.) 68–78 (N.Y. Acad. Sci., New York, 1989).
49. Clayton, D. D. in *14th Texas Symp. on Relativistic Astrophysics* (ed Fenyöes, E. J.) 79–89 (N.Y. Acad. Sci., New York, 1989).
50. Carroll, S. M., Press, W. J. & Turner, E. L. *A. Rev. Astr. Astrophys.* **30**, 499–542 (1992).
51. Pierce, M. J. et al. *Nature* **371**, 385–389 (1994).

ACKNOWLEDGEMENTS. We sincerely thank the many scientists, engineers and astronauts who contributed to HST, making these observations possible. This research was supported by NASA, through a grant from the Space Telescope Science Institute (which is operated by the Association of Universities for Research in Astronomy), and by the NSF. This work has benefitted from the use of the NASA/IPAC Extragalactic Database (NED).

Insulin-dependent stimulation of protein synthesis by phosphorylation of a regulator of 5'-cap function

Arnim Pause^{*†}, Graham J. Belsham^{*†}, Anne-Claude Gingras^{*}, Olivier Donzé^{*}, Tai-An Lin[§], John C. Lawrence Jr[§] & Nahum Sonenberg^{*||}

^{*}Department of Biochemistry and McGill Cancer Centre, McGill University, 3655 Drummond Street, Montréal, Québec H3G 1Y6, Canada

[†]BBSRC Institute for Animal Health, Pirbright, Woking, Surrey GU24 0NF, UK

[§]Department of Molecular Biology and Pharmacology, Washington University School of Medicine, St Louis, Missouri 63110, USA

The cloning is described of two related human complementary DNAs encoding polypeptides that interact specifically with the translation initiation factor eIF-4E, which binds to the messenger RNA 5'-cap structure. Interaction of these proteins with eIF-4E inhibits translation but treatment of cells with insulin causes one of them to become hyperphosphorylated and dissociate from eIF-4E, thereby relieving the translational inhibition. The action of this new regulator of protein synthesis is therefore modulated by insulin, which acts to stimulate the overall rate of translation and promote cell growth.

ALL eukaryotic mRNAs (except organellar) bear a 'cap' structure (m⁷GpppN, where N is any nucleotide and m is a methyl group) at their 5' terminus¹, which is recognized by a cap-binding protein, eIF-4E (ref. 2). This polypeptide exists as a subunit of a complex, termed eIF-4F, which consists of eIF-4E, eIF-4A (a bidirectional RNA helicase) and p220 (for review, see ref. 3). Interaction of the cap-binding complex with the mRNA, followed by unwinding of the mRNA 5' secondary structure, is

thought to facilitate attachment of the smaller 40S ribosomal subunit to mRNA (see ref. 4 for review). The process of initiation of protein synthesis is the rate-limiting step in translation and the point of extensive regulation. Phosphorylation of initiation factors is important in the regulation of translation initiation (reviewed in ref. 5). Addition of the peptide hormone insulin to adipocytes or muscle cells stimulates protein synthesis at the initiation level^{6,8}, and concomitantly increases eIF-4E phosphorylation^{9,10}. This modification enhances the activity of eIF-4E (ref. 11) and partly explains the increased translation rates after treatment of cells with insulin. eIF-4E plays a central

[†] Present address: CBMB, NICHD, NIH, Bethesda, Maryland 20892, USA

^{||} To whom correspondence should be addressed.

FIG. 1 Predicted amino-acid sequence and alignment of 4E-binding proteins. *a*, Sequence alignment of 4E-BP1 and 4E-BP2. *b*, Sequence alignment between human 4E-BP1 and rat PHAS-I.

METHODS. A fusion protein of heart muscle kinase (HMK) phosphorylation site with eIF-4E was expressed in *E. coli* BL 21 cells and purified using a m⁷GDP affinity column as described²⁰. The fusion protein was phosphorylated using [γ -³²P]ATP and bovine HMK (Sigma) to a specific activity of 1×10^8 c.p.m. μ g⁻¹, and used to screen a human placenta λ gt11 cDNA library as described¹⁹. The cDNA inserts of positive λ plaques were released with *Sall*, subcloned into pBluescript KS (Stratagene), and sequenced in both orientations²⁹.

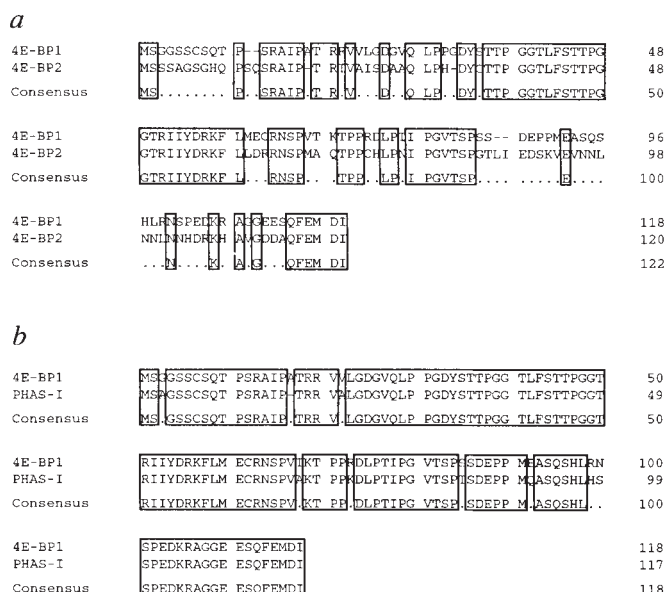
role in control of cell growth, and its inappropriate expression causes major changes in cellular metabolism, as overexpression of eIF-4E results in cell transformation¹². Here we report on the isolation and characterization of a protein whose binding to eIF-4E is regulated by phosphorylation. This protein has been identified as a target for insulin-stimulated phosphorylation in adipose tissue^{13,16} and is a substrate for MAP kinase both *in vivo* and *in vitro*¹⁷. A cDNA encoding this protein, termed PHAS-I, was recently cloned and sequenced¹⁸.

Cloning of cDNAs

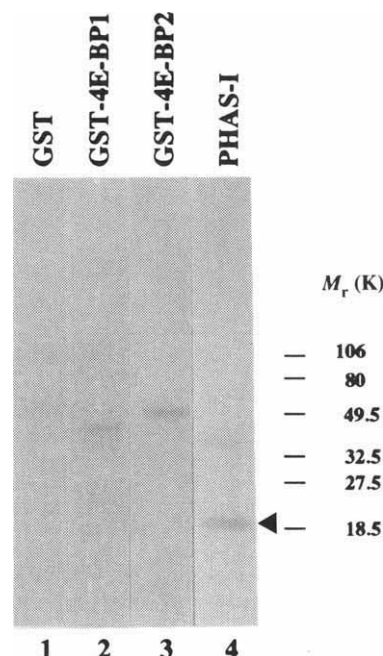
The interaction cloning technique was used to identify proteins that interact with eIF-4E. The eIF-4E cDNA clone was modified at the amino terminus by inserting the phosphorylation site for the heart muscle kinase (HMK)¹⁹. The recombinant protein was expressed in *Escherichia coli* and purified on a m⁷GDP-affinity column to homogeneity²⁰. HMK-eIF-4E was labelled to high specific activity and used in a 'far-western' assay¹⁹, against a λ gt11 cDNA expression library from human placenta. About

FIG. 2 Far-western analysis of the interaction between eIF-4E and 4E-binding proteins or PHAS-I. Purified samples of GST (400 ng), GST-4E-BP1 (600 ng), GST-4E-BP2 (600 ng) and His-tag-PHAS-I (200 ng) were analysed for binding to ³²P-labelled HMK-eIF-4E. Proteins were separated on an SDS-15%-polyacrylamide gel and transferred to nitrocellulose membrane, then incubated with ³²P-labelled HMK-eIF-4E probe (2.5×10^5 c.p.m. ml⁻¹) in 20 ml overnight at 4 °C as described¹⁹ and autoradiographed.

METHODS. GST and GST-fusion proteins were expressed in *E. coli* BL21 and purified as described³⁰. Plasmids were pGEX-2T (containing GST-cDNA), pGEX-2T-4E-BP1, or pGEX-2T[128/129]-4E-BP2. His-tag-PHAS-I was expressed in *E. coli* BL21 (DE3) and purified on a HisBind resin (Novagen). For plasmid pGEX-2T-4E-BP1, 4E-BP1 cDNA was released by restriction-enzyme digestion of pKS-4E-BP1 with *PvuII* and *EcoRI* (this results in the deletion of three N-terminal amino-acids) and subcloned into pGEX-2T restricted with *SmaI* and *EcoRI*. For plasmid pGEX-2T-4EBP2, 4E-BP2 cDNA was released from λ phage DNA by restriction with *EcoRI* and subcloned into plasmid pGEX-2T[128/129], which had been restricted with *EcoRI* (this plasmid contains an HMK flag¹⁹).



one million plaques were screened and nine plaques that interacted with the eIF-4E probe were isolated. The cDNA inserts in the phage were sequenced. Three of the positive plaques contained clones corresponding to the p220-subunit of the cap-binding complex (eIF-4F), as expected. The other positive isolates contained two distinct but related sequences named 4E-BP1 and 4E-BP2 (for 4E-binding proteins 1 and 2; accession numbers for the DNA sequences are L36055 and L36056, respectively). Two phage isolates contained the 4E-BP1 sequence and one isolation of 4E-BP2 was obtained. The cDNAs for 4E-BP1 and 4E-BP2 encode proteins of 118 and 120 amino acids, respectively (Fig. 1). Northern analysis showed, however, that the mRNAs encoding the two proteins are quite different in size: the mRNA encoding 4E-BP1 is about 850 nucleotides long, whereas that of 4E-BP2 is about 2,800 nucleotides (data not shown; the complete sequence of 4E-BP2 mRNA has not yet been determined). The homology between the predicted amino-acid sequences of the two proteins is shown in Fig. 1*a*; the two protein sequences are 56% identical.



eIF-4E-binding proteins

A DNA database search revealed a high degree of identity (93 and 97% similarity) between the protein sequence of human 4E-BP1 and a cDNA encoding a rat protein termed PHAS-I (ref. 18; Fig. 1b), indicating that 4E-BP1 is the human homologue of PHAS-I. The PHAS-I protein is unusual in that it is soluble after boiling and in 2% trichloroacetic acid^{13, 15, 21}. It was initially identified as a protein phosphorylated in response to insulin in rat adipose tissue and in murine 3T3-L1 adipocytes^{13, 16, 21}. Also, it is phosphorylated in response to growth factors and phorbol esters^{21, 22}.

To confirm that 4E-BP1 and 4E-BP2, as well as rat PHAS-I, interact with eIF-4E, recombinant glutathione *S*-transferase fusion proteins (GST-4E-BP1, GST-4E-BP2) and His-tagged PHAS-I were analysed for interaction with eIF-4E by far-western blotting. After SDS-PAGE, purified proteins were transferred to nitrocellulose and probed with ³²P-labelled HMK-eIF-4E. The HMK-eIF-4E probe did not interact with GST alone (Fig. 2, lane 1), but reacted equally with GST-4E-BP1 and GST-4E-BP2 (lanes 2 and 3; GST-4E-BP2 migrates more slowly than GST-4E-BP1 because it contains the HMK sequence). eIF-4E also bound to rat PHAS-I (lane 4, PHAS-I, indicated by a triangle; the slower polypeptide, presumably an aggregate, appears after storage). PHAS-I migrated at a position corresponding to *M_r* 22,000 (22K) on SDS-PAGE, although the size of the protein calculated from its sequence is only ~12K. This also happens with 4E-BPs synthesized *in vitro* (Fig. 3), in agreement with earlier results¹⁸.

We confirmed that 4E-BP1 interacts with eIF-4E by co-immunoprecipitation (using only 4E-BP1[PHAS-I] as no antibody is available against 4E-BP2). RNA transcripts generated *in vitro* were translated together with eIF-4E mRNA in a wheat-

germ translation extract (wheat-germ rather than reticulocyte lysate was used as the latter contained endogenous 4E-BP1 (4E-BP1:eIF-4E ~ 1:20), whereas 4E-BP1 did not bind to endogenous wheat-germ eIF-4E; data not shown) and the translation products immunoprecipitated with anti-PHAS-I antiserum. This antiserum was specific for 4E-BP1 because it did not precipitate eIF-4E synthesized *in vitro* (Fig. 3a; compare lanes 4 and 1). Translation of 4E-BP1 in the wheat-germ extract yielded a polypeptide of ~20K on SDS-PAGE (lane 3), as shown for the rat recombinant protein. This translation product was immunoprecipitated by anti-PHAS-I antiserum (compare lanes 6 and 3). PHAS-I antibody immunoprecipitated eIF-4E and 4E-BP1 together following co-translation of their mRNAs (compare lanes 5 and 2). Similar results were obtained using anti-eIF-4E antiserum; neither protein was immunoprecipitated by preimmune serum (data not shown).

We also demonstrated an interaction between 4E-BP1 and eIF-4E by synthesizing them in wheat-germ translation extract and then incubating them with a m⁷GDP-coupled agarose resin²⁰: eIF-4E, which recognizes the cap structure, bound to the resin, as expected (Fig. 3b, lane 1), but 4E-BP1 did not (lane 2). However, 4E-BP1 was retained by the resin in the presence of eIF-4E (lane 3). This shows that the interaction of 4E-BP1 with eIF-4E does not preclude the binding of eIF-4E to the cap structure.

Insulin and affinity of 4E-BP1 for eIF-4E

As insulin triggers the phosphorylation of PHAS-I (4E-BP1), we tested whether this phosphorylation could modulate the interaction of 4E-BP1 with eIF-4E. Rat adipose tissue was incubated with or without insulin and extracts were heat-treated to enrich for the heat-stable 4E-BP1 protein¹⁴. Soluble proteins

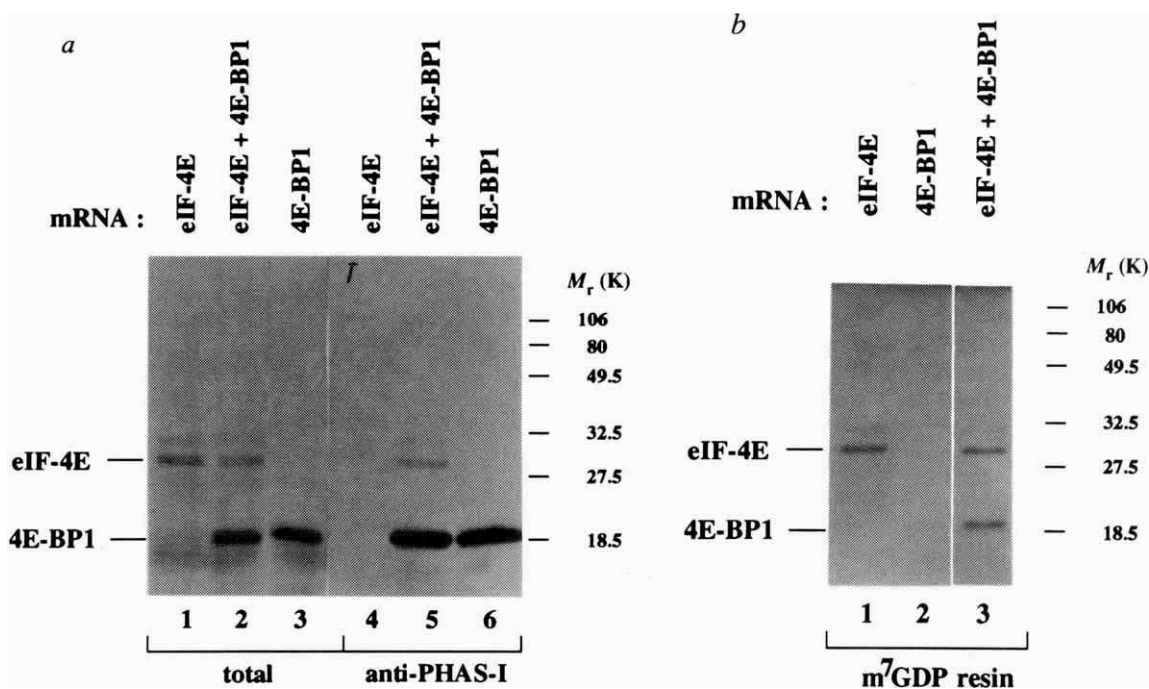


FIG. 3 Complex formation *in vitro* between eIF-4E-BP1. *In vitro* translation products of eIF-4E and 4E-BP1 were analysed by immunoprecipitation with anti-PHAS-I antiserum (a) or by binding to a m⁷GDP (cap) affinity resin (b).

METHODS. Plasmids pKS-eIF-4E and pKS-4E-BP1 were linearized with *Sall* and *Bam*HI respectively, transcribed with T7 and T3 RNA polymerase, respectively, in the presence of m⁷GpppG to yield capped transcripts as described³¹. For a, mRNAs (10 µg ml⁻¹) encoding eIF-4E, 4E-BP1 or both were translated using wheat-germ extract (Promega) with ³⁵S-methionine for 1 h at 25 °C. Samples were then analysed directly or incubated with rabbit anti-PHAS-I antiserum (2 µl) and 300 µl

3% protein A-Sepharose (Pharmacia) suspension in Tris-buffered saline for 1 h at 4 °C, the resin was washed (3 × 1 ml) with 50 mM Tris-HCl, pH 7.5, 150 mM KCl, 1% NP-40, 1 mM EDTA, and eluted proteins were separated on a SDS-15% polyacrylamide gel, fixed and treated with EN³Hance (DuPont), and exposed to X-ray film. For b, mRNAs (4 µg ml⁻¹) encoding eIF-4E (lane 1), 4E-BP1 (lane 2) or both (lane 3) were translated as in a. Translation products were incubated with m⁷GDP affinity resin in buffer A (20 mM Tris-HCl, pH 7.5, 100 mM KCl, 2 mM DTT, 2 mM EDTA), the resin was washed (3 × 1 ml) in buffer A, and bound proteins were eluted in SDS sample buffer and analysed as in a.

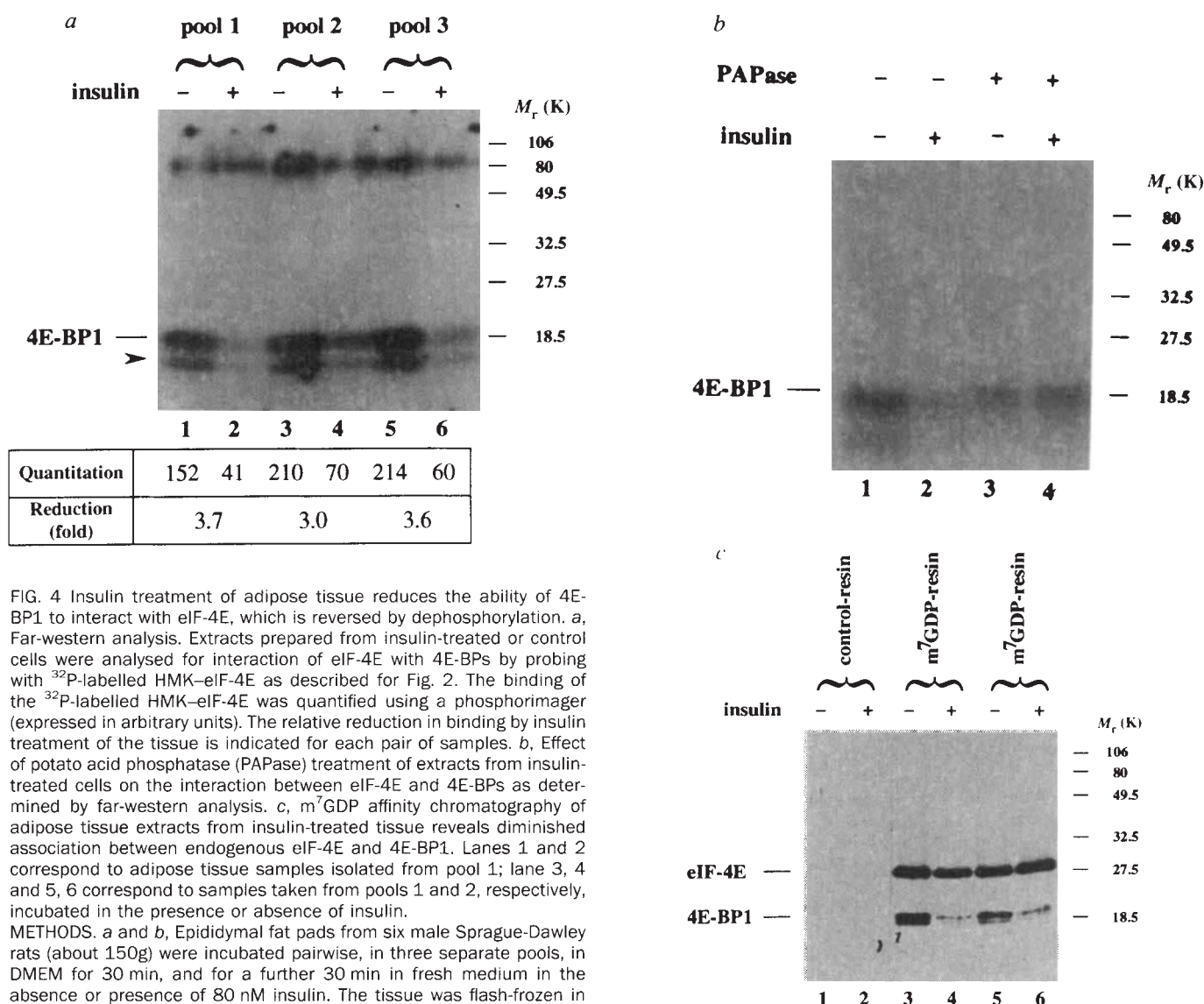


FIG. 4 Insulin treatment of adipose tissue reduces the ability of 4E-BP1 to interact with eIF-4E, which is reversed by dephosphorylation. **a**, Far-western analysis. Extracts prepared from insulin-treated or control cells were analysed for interaction of eIF-4E with 4E-BPs by probing with ³²P-labelled HMK-eIF-4E as described for Fig. 2. The binding of the ³²P-labelled HMK-eIF-4E was quantified using a phosphorimager (expressed in arbitrary units). The relative reduction in binding by insulin treatment of the tissue is indicated for each pair of samples. **b**, Effect of potato acid phosphatase (PAPase) treatment of extracts from insulin-treated cells on the interaction between eIF-4E and 4E-BPs as determined by far-western analysis. **c**, m⁷GDP affinity chromatography of adipose tissue extracts from insulin-treated tissue reveals diminished association between endogenous eIF-4E and 4E-BP1. Lanes 1 and 2 correspond to adipose tissue samples isolated from pool 1; lane 3, 4 and 5, 6 correspond to samples taken from pools 1 and 2, respectively, incubated in the presence or absence of insulin.

METHODS. **a** and **b**, Epididymal fat pads from six male Sprague-Dawley rats (about 150g) were incubated pairwise, in three separate pools, in DMEM for 30 min, and for a further 30 min in fresh medium in the absence or presence of 80 nM insulin. The tissue was flash-frozen in liquid nitrogen and extracted in 3 ml cold buffer containing 0.25 M sucrose, 10 mM Tris-HCl, pH 7.5, 5 mM EDTA, 2 mM EGTA, using a Brinkman PT-10 polytron. The homogenate was centrifuged at 8,000 r.p.m. for 5 min, and the infranant collected. Aliquots were treated at 100 °C for 7 min, the precipitated protein removed by centrifugation, and the remaining protein precipitated with trichloroacetic acid (15%). Samples were analysed by SDS-PAGE, transferred to nitrocellulose and processed for far-western analysis as for Fig. 2. For **b**, extracts were analysed directly or following incubation with potato acid phosphatase (6U) in 100 mM MES, pH 6.0, for 30 min at 37 °C, and repetition of the heat treatment. **c**, Rat adipose tissue was incubated in the absence or

presence of insulin (80 nM) for 30 min and flash-frozen. Tissue extracts were prepared and the infranant was incubated with m⁷GDP-coupled agarose (cap column) (lanes 3–6) or unmodified agarose (lanes 1 and 2), washed (3 × 1 ml) with buffer A (Fig. 3 legend), and bound proteins eluted with SDS sample buffer and analysed by SDS-PAGE. The proteins were transferred to nitrocellulose, incubated in TBS/5% milk/0.1% Tween-20 and then simultaneously with rabbit anti-eIF-4E (1:2,000) and rabbit anti-PHAS-I antibody (2 µg ml⁻¹), followed by horseradish-peroxidase-conjugated anti-rabbit IgG and detected using an ECL kit (Amersham).

were separated by SDS-PAGE, transferred to nitrocellulose membrane and probed with ³²P-labelled HMK-eIF-4E. Two polypeptides migrating at ~20K interact with eIF-4E in extracts from adipose tissue in the absence of insulin. The slower-migrating polypeptide probably corresponds to 4E-BP1, whereas the faster species (indicated by an arrow) may be 4E-BP2, as the protein synthesized *in vitro* in a reticulocyte lysate migrates faster than 4E-BP1 (data not shown; an additional unknown polypeptide of M_r ~85K also interacted with eIF-4E). Treatment with insulin led to a marked reduction (about 3–4-fold in three separate pools of adipose tissue) in the interaction of the two ~20K polypeptides with eIF-4E (Fig. 4a; compare lanes 2 and 4, 6 and 1, 3 and 5). Analysis of the same extracts using an antiserum raised against the C terminus of PHAS-I showed that treatment with insulin did not alter the total amount of 4E-BP1 (data not

shown). These results demonstrate that the interaction between exogenous eIF-4E and 4E-BP1 is decreased in response to insulin.

To show directly that phosphorylation of 4E-BP1 upon insulin treatment is responsible for the dissociation of 4E-BP1 from eIF-4E, extracts were treated with acid phosphatase. Insulin treatment resulted in decreased interaction between eIF-4E and 4E-BP1 (Fig. 4b; compare lanes 2 and 1). But when samples were treated with phosphatase, the binding of eIF-4E to 4E-BP1 in insulin-treated extracts was increased to a level comparable to control extracts (compare lanes 4 and 3; the amount of protein recovered after phosphatase treatment and re-boiling was reduced relative to untreated extracts, which explains the weaker signal in lane 3 compared to lane 1; number of experiments, $n=3$).

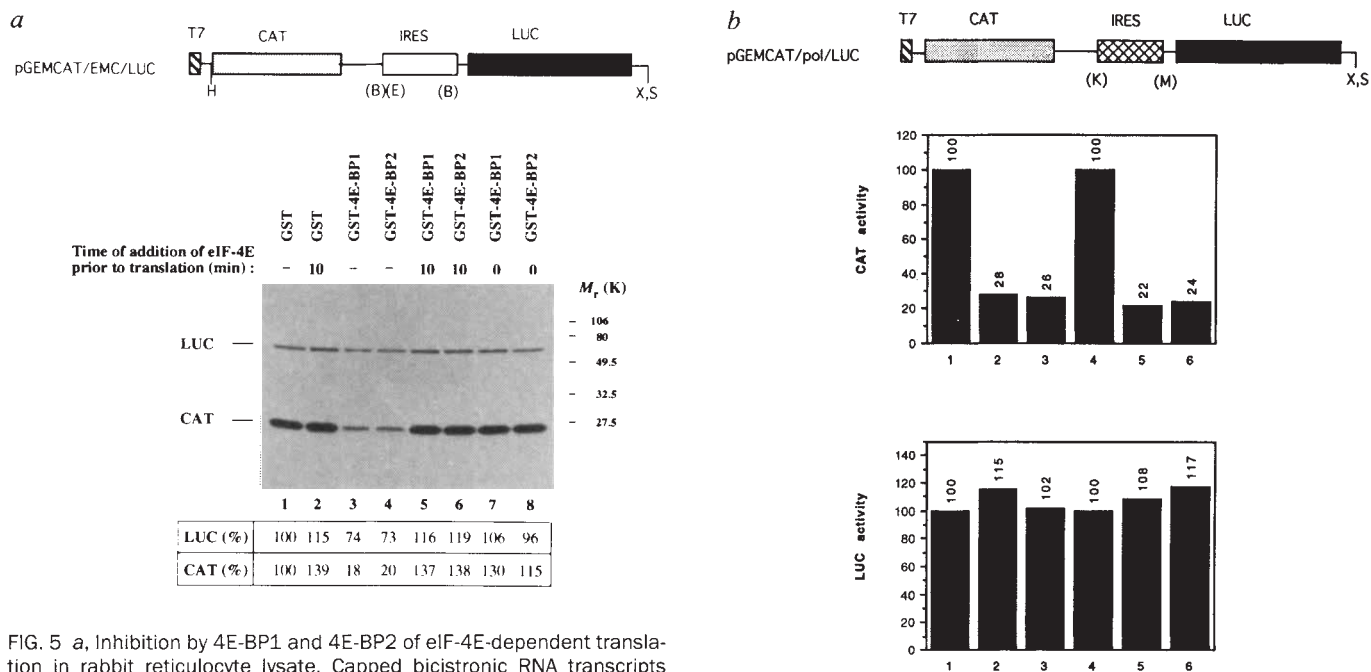


FIG. 5 **a**, Inhibition by 4E-BP1 and 4E-BP2 of eIF-4E-dependent translation in rabbit reticulocyte lysate. Capped bicistronic RNA transcripts encoding both chloramphenicol acetyltransferase (CAT) and luciferase (LUC), were translated in rabbit reticulocyte lysate (Promega). The lysate was preincubated for 10 min with purified GST (400 ng) (lanes 1, 2), GST-4E-BP1 (600 ng) (lanes 3, 5, 7) or GST-4E-BP2 (600 ng) (lanes 4, 6, 8). Recombinant eIF-4E (400 ng) was added where indicated, either at the beginning of the preincubation (lanes 7, 8) or at the start of translation (lanes 2, 5, 6). **b**, Inhibition of cap-dependent protein synthesis *in vivo* by expression of 4E-binding proteins.

METHODS. **a**, RNA was produced using T7 RNA polymerase from *Xho*I-linearized pGEMCAT/EMC/LUC (top). The IRES (internal ribosome entry site) is derived from the 5' untranslated region of encephalomyocarditis virus (EMCV). The indicated restriction enzyme sites are as follows: H, *Hind*III; B, *Bam*HI; E, *Eco*RI; S, *Sac*I, X, *Xho*I. Translation was carried out according to the manufacturer's instructions with $7 \mu\text{g ml}^{-1}$ mRNA in $15 \mu\text{l}$ for 30 °C before analysis by SDS-PAGE. Translation of CAT and LUC was quantified using a phosphorimager, and expressed as arbitrary units. Incorporation into CAT and LUC in the presence of GST alone (lane 1) was set at 100%. **b**, Human TK⁻ 143B cells (ATCC CRL 8303)

were infected with the recombinant vaccinia virus vTF7-3, which expresses T7 RNA polymerase. Cells were transfected with plasmids pGEMCAT/pol/LUC (top) (2 μg ; lanes 1–3) or pGEMCAT/EMC/LUC (2 μg ; lanes 4–6), encoding bicistronic mRNAs (top). pGEM/pol/LUC contains the poliovirus type-II Lansing IRES. Restriction enzyme sites as in **a**; K, *Kpn*I, and M, *Msc*I. Plasmids (5 μg) derived from pRC/CMV (Invitrogen) and containing the coding sequence for 4E-BP1 (lanes 2, 5) or 4E-BP2 (lanes 3, 6) under control of the T7 promoter were cotransfected with the reporter constructs using lipofectin (15 μg) (BRL). After 20 h, cell extracts were prepared and analysed for CAT expression by enzyme-linked immunosorbent assay (Boehringer) and LUC activity was measured in millivolts in a luminometer (BIOORBIT) using a luciferase assay system (Promega). Expression of each reporter protein assayed from the reporter vector alone was set at 100% in each case. The experiment was carried out twice, with <10% variation between the experiments.

To determine whether insulin affects the interaction of 4E-BP1 with eIF-4E *in vivo*, we used fat-cell extracts from untreated and insulin-treated tissue to assay for the ability of 4E-BP1 to bind through eIF-4E to a $m^7\text{GDP}$ affinity column. Bound proteins were eluted with SDS-sample buffer, resolved by SDS-PAGE, then immunoblotted with anti-PHAS-I and anti-eIF-4E antibodies. Neither eIF-4E nor PHAS-I bound to the resin not containing the cap (Fig. 4c, lanes 1 and 2). Both eIF-4E and PHAS-I bound to the $m^7\text{GDP}$ resin when extracts from untreated control cells were used (two pools of tissue were analysed; lanes 3 and 5). But although the amount of PHAS-I from insulin-treated cells that bound to the cap column was significantly reduced, similar levels of eIF-4E were bound in each case (lanes 3–6). Consistent with these results, in 3T3-L1 cells incubated with $^{32}\text{P}_i$, PHAS-I antiserum immunoprecipitated ^{32}P -labelled PHAS-I and eIF-4E. Insulin treatment enhanced the phosphorylation of PHAS-I and reduced the amount of eIF-4E precipitating with PHAS-I (data not shown). We therefore conclude that the phosphorylation of 4E-BP1 in response to insulin exposure results in 4E-BP1 dissociation from eIF-4E, indicating that this may account for the enhanced translation activity in adipose tissue upon insulin treatment.

Inhibition of cap-dependent translation

To investigate the possibility that interaction of the 4E-BPs with eIF-4E interferes with cap-dependent but not cap-independent

translation, we made use of artificial bicistronic mRNAs containing the internal ribosome entry site (IRES) from encephalomyocarditis virus (EMCV) (ref. 23; Fig. 5a). The translation of chloramphenicol acetyltransferase (CAT) is cap- and eIF-4E-dependent, but the translation of luciferase (LUC) directed by the EMCV IRES is cap-independent²³. Translation of this RNA transcript in rabbit reticulocyte lysate produced both CAT and luciferase efficiently (Fig. 5a). Preincubation of the lysate for 10 min before addition of RNA with purified GST-4E-BP1 or GST-4E-BP2 specifically inhibited translation of CAT by ~5-fold (lanes 3 and 4, respectively). This inhibition could be relieved by addition of recombinant eIF-4E, regardless of whether it was present during (lanes 5 and 6), or after (lanes 7 and 8) the preincubation of 4E-BPs with reticulocyte lysate. This suggests that the formation of the 4E-BP-eIF-4E complex does not irreversibly inhibit the translational machinery. In contrast to the deleterious effect of 4E-BPs on cap-dependent translation, there was no significant effect (1.3-fold inhibition) on IRES-directed luciferase translation following addition of GST-4E-BPs (compare lanes 3 and 4 to lane 1).

To investigate whether 4E-BPs inhibit cap-dependent translation *in vivo*, the same bicistronic reporter system was analysed using a transient expression assay in cells infected with the recombinant vaccinia virus vTF7-3, which expresses T7 RNA polymerase^{24,25}. Expression of CAT and luciferase was efficient from pGEM-CAT/pol/LUC or pGEM-CAT/EMC/LUC

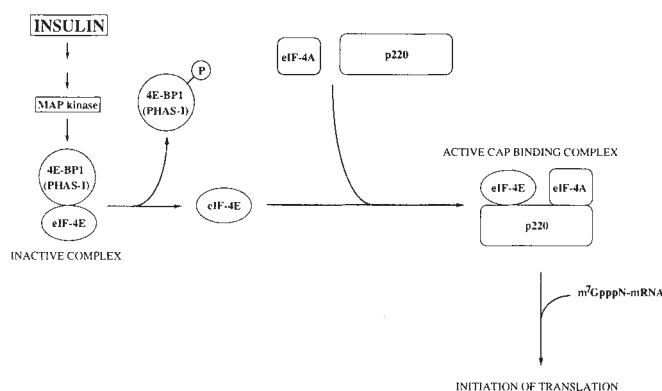


FIG. 6 A model for the mechanism of stimulation of translation by insulin. The model presumes that binding of 4E-BP1 to eIF-4E prevents the interaction of the latter with p220. Insulin-dependent phosphorylation of 4E-BP1 by MAP kinase releases eIF-4E, and makes it available to form an active cap-binding complex (eIF-4F). Although not indicated in the model, insulin stimulates the phosphorylation of eIF-4E, possibly through a protein kinase C-dependent pathway⁹, which is also likely to enhance translation initiation.

mRNAs when they were assayed alone. When plasmids encoding 4E-BP1 or 4E-BP2 were cotransfected with either of the reporter plasmids, there was specific inhibition of CAT expression (4–5-fold) but little effect on luciferase expression (Fig. 5b). Hence, 4E-BP1 and 4E-BP2 are specific inhibitors of cap-dependent translation both *in vivo* and *in vitro*. These results also demonstrate that the requirement for eIF-4E differs for cap-dependent and IRES-directed translation.

Discussion

We have isolated cDNA clones from a human placenta cDNA library encoding two previously unknown eIF-4E-binding proteins. The sequence of human 4E-BP1 is almost identical to that of PHAS-I from rat adipose tissue and skeletal muscle. In adipose tissue stimulated by insulin, PHAS-I is one of the major phosphorylation targets, together with ribosomal protein S6 and ATP citrate lyase¹⁶. We have shown that purified PHAS-I has eIF-4E-binding activity. The insulin-dependent increase in the phosphorylation of 4E-BP1 (PHAS-I) in adipose tissue dramatically diminishes the ability of the protein to interact with eIF-4E. Interaction of 4E-BP1 with eIF-4E inhibits cap-dependent protein synthesis both *in vitro* and *in vivo*, but has only a small effect on cap-independent protein synthesis. Thus, 4E-BP1 must be specifically affecting the process of cap-dependent translation initiation. The data support a model (Fig. 6), in which the phosphorylation of 4E-BP1 by MAP kinase, in response to insulin treatment, results in its dissociation from eIF-4E. This presum-

ably allows eIF-4E to interact with the p220 subunit to form a functional eIF-4F complex, which is required for translation initiation^{1–4}. The finding that PHAS-I (4E-BP1) is phosphorylated by MAP kinase both *in vivo* and *in vitro*¹⁷ is highly significant as it directly links the insulin signal transduction pathway to translation initiation. Insulin treatment of 3T3-L1 cells, or NIH 3T3 cells overexpressing the insulin receptor also leads to an increase in the phosphorylation of eIF-4E (refs 9, 10), and phosphorylation of eIF-4E is thought to be required for its activity¹¹. The enhanced phosphorylation of eIF-4E in insulin-treated tissue may therefore also contribute to the decreased interaction between eIF-4E and 4E-BPs. But phosphorylation of 4E-BP1 alone is sufficient to modulate the interaction, because the eIF-4E probe used to detect the interaction was made in *E. coli* and was not phosphorylated at an authentic site.

Our model provides an explanation for data demonstrating that insulin enhances translation initiation in adipocytes and in muscle cells^{6–8} about 2–3-fold. Insulin specifically stimulates the translation of capped, but not uncapped, mRNAs in 3T3-L1 cells²⁶, paralleling the discriminatory inhibition of 4E-BPs on cap-dependent translation: these experiments involved transfection of capped or uncapped luciferase mRNAs into 3T3-L1 cells followed by insulin treatment, suggesting that insulin enhances the function of the cap-binding complex²⁶. The changes in the binding activity of 4E-BP1 described here would explain these observations.

Because two closely related proteins, each capable of efficiently binding eIF-4E, could be isolated from the placental cDNA library, it is possible that there is a family of such related proteins in different cell types or even within the same cell. Different tissues gave only a weak northern blot signal with a PHAS-I cDNA probe, except for muscle and adipose tissue, for which the signal was strong¹⁸. Inasmuch as this analysis would not be expected to detect the related 4E-BP2 mRNA, it could be present in other tissues, so the two (and possibly more) members of the 4E-BP family could be expressed differentially in a tissue-specific manner.

Overexpression of eIF-4E results in cell transformation¹². It has been proposed that overexpression of eIF-4E increases the amount of eIF-4F complex and enhances translation initiation of mRNAs encoding growth-promoting proteins. Indeed, translation of cyclin D1 (ref. 27) and ornithine decarboxylase is stimulated in cells overexpressing eIF-4E (ref. 28). Also, insulin preferentially stimulates translation of ornithine decarboxylase transcripts in fibroblasts overexpressing the insulin receptor¹⁰. Our findings are consistent with this model. Hyperphosphorylation of 4E-BP1 (PHAS-I), which enhances eIF-4E activity, also occurs in response to growth-promoting factors like EGF and PDGF²¹, which activate the MAP kinase pathway. 4E-BP1, 4E-BP2 and related proteins may therefore also play a key role in regulating translation rates and cell growth in response to treatment with growth-promoting factors. □

Received 11 July; accepted 3 October 1994.

- Shatkin, A. J. *Cell* **40**, 223–224 (1985).
- Merrick, W. C. *Microbiol. Rev.* **56**, 291–315 (1992).
- Sonenberg, N. *Gene Expression* **3**, 317–323 (1993).
- Hershey, J. W. B. *Rev. Biochem.* **60**, 717–757 (1991).
- Hershey, J. W. B. *J. Biol. Chem.* **264**, 20823–20826 (1989).
- Lyons, R. T., Nordeen, S. K. & Young, D. A. *J. Biol. Chem.* **255**, 6330–6334 (1980).
- Morgan, H. E. et al. *J. Biol. Chem.* **246**, 2163–2170 (1971).
- Stirwalt, W. S., Wool, I. G. & Cavicchi, P. *Proc. natn. Acad. Sci. U.S.A.* **57**, 1885–1889 (1967).
- Morley, S. J. & Traugh, J. A. *J. Biol. Chem.* **265**, 10611–10616 (1990).
- Manzella, J. M., Rychlik, W., Rhoads, R. E., Hershey, J. W. B. & Blackshear, P. J. *J. Biol. Chem.* **266**, 2383–2389 (1991).
- Joshi-Barve, S., Rychlik, W. & Rhoads, R. E. *J. Biol. Chem.* **265**, 2979–2983 (1990).
- Lazaris-Karatzas, A., Montine, K. S. & Sonenberg, N. *Nature* **345**, 544–547 (1990).
- Belsham, G. J., Brownsey, R. W., Hughes, W. A. & Denton, R. M. *Diabetologia* **18**, 307–312 (1980).
- Belsham, G. J., Brownsey, R. W. & Denton, R. M. *Biochem. J.* **204**, 345–352 (1982).
- Denton, R. M., Brownsey, R. W. & Belsham, G. J. *Diabetologia* **21**, 347–363 (1981).
- Avruch, J. et al. in *Molecular Basis of Insulin Action* (ed. Czech, M. P.) 67–94 (Plenum, New York, 1985).
- Haystead, T. A. J. et al. *J. Biol. Chem.* **269**, 23185–23191 (1994).
- Hu, C., Pang, S., Kong, X., Velleca, M. & Lawrence, J. C. *Proc. natn. Acad. Sci. U.S.A.* **91**, 3730–3734 (1994).

- Blancar, M. A. & Rutter, W. J. *Science* **256**, 1014–1018 (1992).
- Edery, I., Altmann, M. & Sonenberg, N. *Gene* **74**, 517–525 (1988).
- Blackshear, P. J., Nemenoff, R. A. & Avruch, J. *Biochem. J.* **214**, 11–19 (1983).
- Blackshear, P. J. et al. *J. Biol. Chem.* **260**, 13304–13315 (1985).
- Jackson, R. J., Howell, M. T. & Kaminski, A. *Trends biochem. Sci.* **15**, 477–483 (1990).
- Fuerst, T. R., Niles, E. G., Studier, F. W. & Moss, B. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8122–8126 (1986).
- Belsham, G. J. & Brangwyn, J. K. *J. Virol.* **64**, 5389–5395 (1990).
- Gallie, D. R. & Traugh, J. A. *J. Biol. Chem.* **269**, 7174–7179 (1994).
- Rosenwald, I. B., Lazaris-Karatzas, A., Sonenberg, N. & Schmidt, E. V. *Molec. cell. Biol.* **13**, 7358–7363 (1993).
- Shantz, L. M. & Pegg, A. E. *Cancer Res.* **54**, 2313–2316 (1994).
- Hattori, M. & Sakaki, Y. *Analyt. Biochem.* **152**, 232–238 (1986).
- Methot, N., Pause, A., Hershey, J. W. B. & Sonenberg, N. *Molec. cell. Biol.* **14**, 2307–2316 (1994).
- Pelletier, J. & Sonenberg, N. *Cell* **40**, 515–526 (1985).

ACKNOWLEDGEMENTS. We thank C. Lister for help in sequencing the 4E-BP cDNAs, M. Park and M. Blancar for the human placenta cDNA library and for plasmids pAR(ΔR1) and pGEX-2T[128/129], respectively; and J. Pelletier, M. Szyf, N. Methot and H. Yamataka for critical comments on the manuscript. This work was supported by grants from the National Cancer Institute of Canada (N.S.) and from the NIH (J.C.L.). A.P. was supported by a Cancer Research Society of Montreal studentship.

Mitochondrial Hsp70/MIM44 complex facilitates protein import

Hans-Christoph Schneider, Jutta Berthold, Matthias F. Bauer, Klaus Dietmeier*, Bernard Guiard*, Michael Brunner† & Walter Neupert

Institut für Physiologische Chemie der Universität München, Goethestraße 33, 80336 München, Germany

Protein translocation into mitochondria requires the mitochondrial protein Hsp70. This molecular chaperone of the mitochondrial matrix is recruited to the protein import machinery by MIM44, a component associated with the inner membrane of the mitochondria. Formation of the mt-Hsp70/MIM44 complex is regulated by ATP. MIM44 and mt-Hsp70 interact in a sequential manner with incoming segments of unfolded preproteins and thereby facilitate stepwise vectorial translocation of proteins across the mitochondrial membranes. The complex appears to act as a molecular ratchet which is energetically driven by the hydrolysis of ATP.

MITOCHONDRIA contain two translocation systems for the transport of proteins from the cytosol into the matrix, one in the outer membrane and another one in the inner membrane^{1,2}. Both systems cooperate²⁻⁴ but can be functionally and physically separated⁵⁻⁷. The translocation machinery of the outer membrane consists of a complex of at least six protein components⁸.

With the translocation system of the inner membrane three proteins have been described. Mas6p⁹, also termed MIM23¹⁰, is an essential component of the inner membrane. Preproteins were found in contact with this protein during translocation¹¹. MIM17¹⁰, also termed Smlp¹², a component also integrated into the inner membrane, was identified as being essential for protein import into the matrix. The third component, MIM44¹³, also termed Mpi1p¹⁴ or ISP45¹⁵, is a hydrophilic protein. MIM44 is associated with the inner face of the inner membrane¹³ but it is not clear whether its carboxy terminus is exposed to the intermembrane space^{13,16}.

A central role in import has also been assigned to mitochondrial Hsp70 (mt-Hsp70)¹⁷. Preproteins arrested in transit across both mitochondrial membranes were found to be in contact with mt-Hsp70¹⁸⁻²⁰. Mitochondrial Hsp70 was proposed to interact in an ATP-dependent fashion with the incoming unfolded polypeptide chains and thereby to drive energetically the vectorial movement of preproteins into the matrix space^{21,24}.

Here we show that MIM44 plays a central role in preprotein import by cooperating with mt-Hsp70. MIM44 forms a complex with mt-Hsp70 in the presence of ATP. On ATP hydrolysis the complex disintegrates. Both MIM44 and mt-Hsp70 make contact with the polypeptide chain in transit. Mutated mt-Hsp70s, with conditional defects in import of preproteins, have a greatly reduced ability for complex formation under non-permissive conditions. Our results suggest a mechanism of mt-Hsp70/MIM44 function by which MIM44 first binds the unfolded polypeptide chain emerging from the putative translocation channel, then transfers it to mt-Hsp70 concomitant with ATP hydrolysis and complex dissociation. This allows the polypeptide chain to move further into the mitochondria and undergo another cycle of binding to the complex. The pathway proposed here integrates mechanistic aspects of transport such as unfolding and requirements for chaperones with energetic aspects of vectorially driving

polypeptide chain movement across the two mitochondrial membranes.

A complex of MIM44 with mt-Hsp70

To characterize proteins in contact with MIM44 we did immunoprecipitation experiments with affinity-purified MIM44 immunoglobulin G (IgG) using mitochondria solubilized in Triton X-100. In addition to MIM44 a protein with a molecular mass of 70,000 (M_r 70K) and one of 22K were immunoprecipitated (Fig. 1a).

By immunoblotting, the 70K protein was identified as mt-Hsp70 (Fig. 1b) and the 22K protein as MGE (Fig. 1c), the mitochondrial homologue of bacterial GrpE^{25,27}. Mdj1p, the mitochondrial DnaJ homologue²⁸, could not be detected. The apparent molar ratio of mt-Hsp70 to MIM44 in the complex was close to unity as determined by quantification of the radioactivities in the 70K and the 44K bands. MGE was present in significantly lower amounts. When detergent-lysed mitochondria were subjected to gel filtration, MIM44 was recovered in a larger assembly of about 250K (not shown). Thus, the complex appears to consist of two molecules of each, MIM44 and mt-Hsp70.

On treatment of intact mitochondria with the chemical crosslinker disuccinimidyl suberate (DSS) specific MIM44-containing adducts were generated (Fig. 2a). Low concentrations of DSS (25 μ M) were sufficient to form major adducts with MIM44 with apparent M_r s of 116K and 160K. The adducts were extractable with carbonate (pH 11), indicating that MIM44 was not crosslinked to an integral membrane protein. After immunoprecipitation with anti-MIM44 IgG the adducts could be decorated with antibodies against both MIM44 and mt-Hsp70. This indicates that in intact mitochondria MIM44 is in contact with mt-Hsp70. The 116K adduct is apparently composed of one MIM44 and one mt-Hsp70, whereas the 160K adduct could contain two MIM44 and one mt-Hsp70. The formation of a ternary adduct might be due to the heterotetrameric composition of the mt-Hsp70/MIM44 complex and to the excessively high number of lysine residues in MIM44¹⁴.

Mt-Hsp70 and MIM44 make up roughly 1% and 0.1% of total mitochondrial protein, respectively (data not shown). We asked whether, on overexpression of MIM44, mt-Hsp70 would be recruited for complex formation. For this purpose two His-tagged versions of MIM44, MIM44 Δ 3_{His6} and MIM44 Δ 12_{His6} were constructed, in which 3 or 12 C-terminal amino-acid residues were replaced by 6 histidine residues. The overexpressed MIM44 proteins were found to be peripherally associated with

* Present addresses: Biochemisches Institut der Universität Freiburg, Hermann-Herder-Straße 7, 79104 Freiburg, Germany (K.D.); Centre de Genetique Moleculaire CNRS, Universite Pierre et Marie Curie, 91190 Gif-sur-Yvette, France (B.G.).

† To whom correspondence should be addressed.

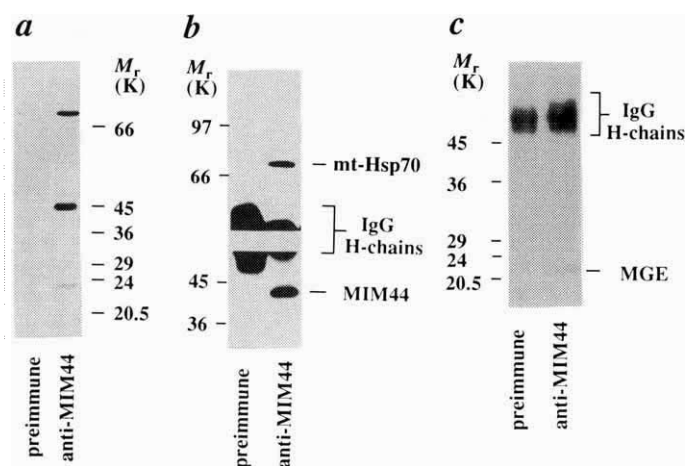


FIG. 1 MIM44 forms a complex with mt-Hsp70 and MGE. **a**, Radiolabelled mitochondria were solubilized in 0.5% Triton X-100 in the presence of EDTA and subjected to immunoprecipitation with anti-MIM44 IgG or preimmune IgG. Immunoprecipitates were analysed by SDS-PAGE and fluorography. **b**, Immunoprecipitates from unlabelled mitochondria were resolved by SDS-PAGE, transferred to nitrocellulose and decorated with anti-mt-Hsp70 IgG (upper part) and with anti-MIM44 IgG (lower part). **c**, Immunoprecipitates were decorated with antibodies against MGE. **METHODS**. Antisera against a His-tagged MIM44 fragment (*EcoRV*-*PstI*) were affinity-purified using the recombinant antigen. Mitochondria were prepared⁴⁷ from yeast cells (D273-10B) labelled with ³⁵S-sulphate. Mitochondria (100 µg) were solubilized in 1 ml IP buffer (0.5% (w/v) Triton X-100, 30 mM Tris-HCl, 100 mM NaCl, 0.5 mM EDTA, 0.5 mM EGTA, 0.5 mM o-phenanthroline, 0.5 mM PMSF, 1.0 mg ml⁻¹ fatty acid-free BSA, and 0.1 mg ml⁻¹ α₂-makroglobulin, pH 7.4). The supernatant of a clarifying spin (30 min, 4 °C, 105,000g in a Beckman TLA45 rotor) was incubated under gentle shaking for 1 h at 4 °C with affinity-purified anti-MIM44 IgG or preimmune serum bound to protein A-sepharose. Goat-anti-rabbit antibodies conjugated to horseradish peroxidase (Sigma) and detection by chemoluminescence (ECL, Amersham) were used for immunodecoration in **b**, and alkaline phosphatase-coupled secondary antibodies in **c**.

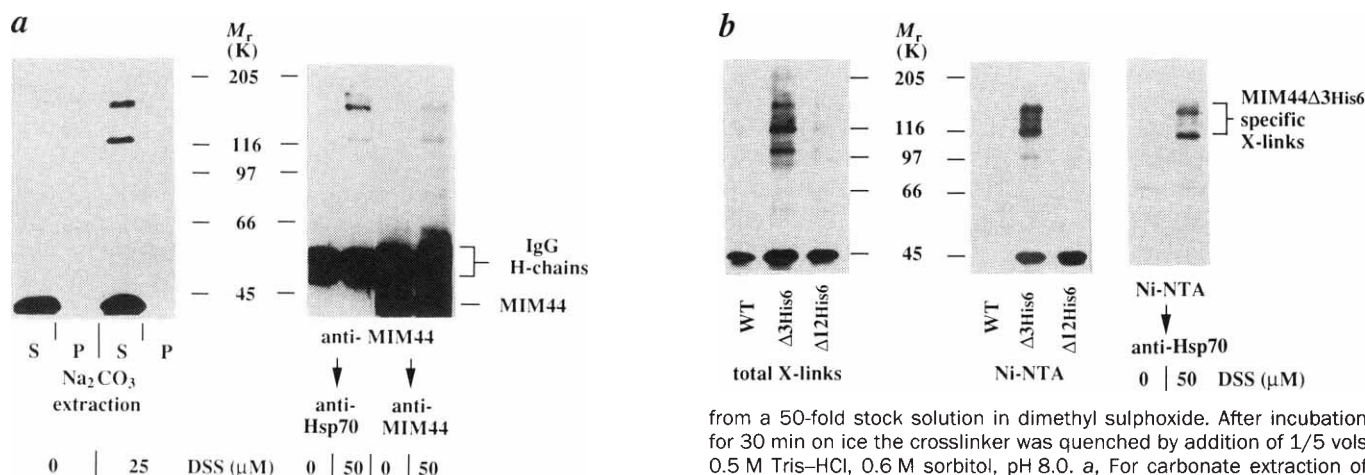


FIG. 2 MIM44 can be crosslinked to mt-Hsp70 in intact mitochondria. **a**, Mitochondria were preincubated with or without DSS and then subjected to extraction with carbonate (pH 11). Supernatants and pellets were analysed by SDS-PAGE and immunoblotting with anti-MIM44 IgG (left panel). MIM44-specific adducts were immunoprecipitated with anti-MIM44 IgG under denaturing conditions. The immunoprecipitate was resolved by SDS-PAGE and decorated with antibodies against mt-Hsp70 and MIM44 (right panel). **b**, MIM44 Δ 3_{His6}, but not MIM44 Δ 12_{His6}, forms crosslinks with mt-Hsp70. Mitochondria harbouring either wild-type MIM44 (WT) or in addition MIM44 Δ 3_{His6} and MIM44 Δ 12_{His6}, respectively, were subjected to crosslinking with 50 µM DSS. Crosslinks were resolved by SDS-PAGE and detected by immunostaining with anti-MIM44 IgG. His-tagged proteins and corresponding cross-linked adducts were isolated under denaturing conditions by Ni-NTA chromatography and analysed by SDS-PAGE and immunostaining with anti-MIM44 IgG. MIM44 Δ 3_{His6}-specific crosslinks were also decorated with anti-mt-Hsp70 IgG.

METHODS. Crosslinking was done in 10 mM HEPES/KOH, 0.6 M sorbitol, pH 8.0 at 1 mg ml⁻¹ mitochondrial protein. DSS (Pierce) was added

from a 50-fold stock solution in dimethyl sulphoxide. After incubation for 30 min on ice the crosslinker was quenched by addition of 1/5 vols 0.5 M Tris-HCl, 0.6 M sorbitol, pH 8.0. **a**, For carbonate extraction of peripheral membrane proteins, mitochondria (10 mg ml⁻¹) were diluted 100-fold in a freshly prepared solution of 0.1 M Na₂CO₃, pH 11.0. The sample was kept on ice for 20 min and centrifuged for 60 min at 120,000g in a Beckman 50Ti rotor. Proteins in the membrane pellet and supernatant (after precipitation with 12% TCA) were analysed by SDS-PAGE. Chemoluminescence (ECL, Amersham) was used for visualization. **b**, For expression of His-tagged MIM44 proteins, Yep51⁴⁸ was digested with *NaeI*-*Sall*, the ends were filled in with Klenow and re-ligated. A truncated MIM44 cDNA fragment was excised from pEP30 (gift from M. Meijer) with *Bam*HI-*Bgl*III and subcloned into pQE-18 (Qiagen). The resulting His-tagged gene, MIM44 Δ 12_{His6}, was excised with *Bam*HI-*Hind*III and introduced into the modified Yep51. The double-stranded synthetic DNA fragment (encoding a His6-tag) 5'-CCGATA-TCTAGACATCACCATCACCATCAGTCGACAAGCTTGCC-3' was digested with *Xba*I and *Hind*III and introduced into the *Xba*I and *Hind*III sites of pEP30. From the resulting plasmid the cDNA encoding MIM44 Δ 3_{His6} was excised with *Bam*HI and *Hind*III and subcloned into the modified Yep51. The resulting Yep51 plasmids were transformed into the yeast strain 334⁴⁹ and transformants were grown on lactate medium in the presence of 1% galactose.

the inner membrane, indicating that a possible component mediating membrane association would not be limiting. MIM44 Δ 3_{His6} was able to complement a MIM44 disruption mutant, whereas MIM44 Δ 12_{His6} was not (not shown). In mitochondria harbouring MIM44 Δ 3_{His6} the yields of the crosslinked products were significantly increased as compared to mitochondria containing only wild-type MIM44 (Fig. 2b). When mito-

chondria containing MIM44 Δ 3_{His6} were subjected to chromatography on nitrilo-tri-acetic acid (Ni-NTA), the His-tagged MIM44 was retained on the resin. In addition, the crosslink products were recovered on Ni-NTA. The 116K and 160K adducts were recognized by antibodies against both mt-Hsp70 and MIM44. The 97K species, recognized only by anti-MIM44 IgG could represent a MIM44 dimer. Neither MIM44

nor mt-Hsp70 specific adducts were recovered on Ni-NTA when mitochondria without His-tagged MIM44 were used. Thus overexpression of MIM44 Δ _{His6} is sufficient to recruit mt-Hsp70 into the complex. MIM44 apparently is the limiting component for complex formation.

With MIM44 Δ _{His6} no significant amounts of crosslinked products were observed on Ni-NTA chromatography. As no crosslinkable lysyl residues were removed by the C-terminal deletion, MIM44 Δ _{His6} does not form a stable complex with mt-Hsp70. This, together with the inability of MIM44 Δ _{His6} to rescue a MIM44 disruption mutant, suggests that complex formation with mt-Hsp70 is crucial for the function of MIM44.

Regulation of Mt-Hsp70/MIM44 interaction

As molecular chaperones of the Hsp70 class possess ATPase activity²⁹, we investigated whether the mt-Hsp70/MIM44 complex is dependent on nucleotides. The sedimentation behaviour of MIM44 was analysed by glycerol gradient centrifugation in the presence of either EDTA or Mg-ATP (Fig. 3a). MIM44 sedimented in a high M_r form when EDTA was present but in a lower M_r form when Mg-ATP was present. This indicates that the mt-Hsp70/MIM44 complex disassembles in the presence of Mg-ATP.

Mitochondria were solubilized in Triton X-100 either in the presence of EDTA or Mg-ATP, and MIM44 was immunoprecipitated with affinity-purified IgGs which were covalently

coupled to Sepharose beads. In the presence of Mg-ATP only MIM44 was detected in the immunoprecipitate whereas after solubilization in EDTA roughly stoichiometric amounts of mt-Hsp70 were coprecipitated (Fig. 3b).

The nucleotide specificity of complex dissociation was investigated using the coimmunoprecipitation assay (Fig. 3c). Lysis of mitochondria in the presence of Mg-ATP resulted in complete dissociation. By contrast, the complex was stable when the non-hydrolysable ATP analogue AMP-PNP was present. Likewise EDTA preserved the complex. When mitochondria were lysed in the presence of Mg²⁺-ions or Mg-ADP the amount of mt-Hsp70 in complex with MIM44 was dependent on the concentration of mitochondria during lysis (not shown). Apparently the low concentration of ATP originating from the endogenous mitochondrial nucleotide pool led to partial dissociation of the complex. Mg-ATP-dependent dissociation of mt-Hsp70 could be competed by Mg-AMP-PNP, but not by Mg²⁺-ions or by Mg-ADP. Thus, hydrolysis of ATP is required for the dissociation of mt-Hsp70 from MIM44.

We then studied the influence of unfolded protein on the stability of the mt-Hsp70/MIM44 complex. Mitochondria were lysed in the presence of synthetic peptides or carboxymethylated lactalbumin (RCMLA), a permanently unfolded protein which is known to bind efficiently to Hsp70s^{30,31}. In the presence of EDTA the amount of mt-Hsp70 complexed with MIM44 could not be diminished by peptides or RCLMA (Fig. 3d). Partial

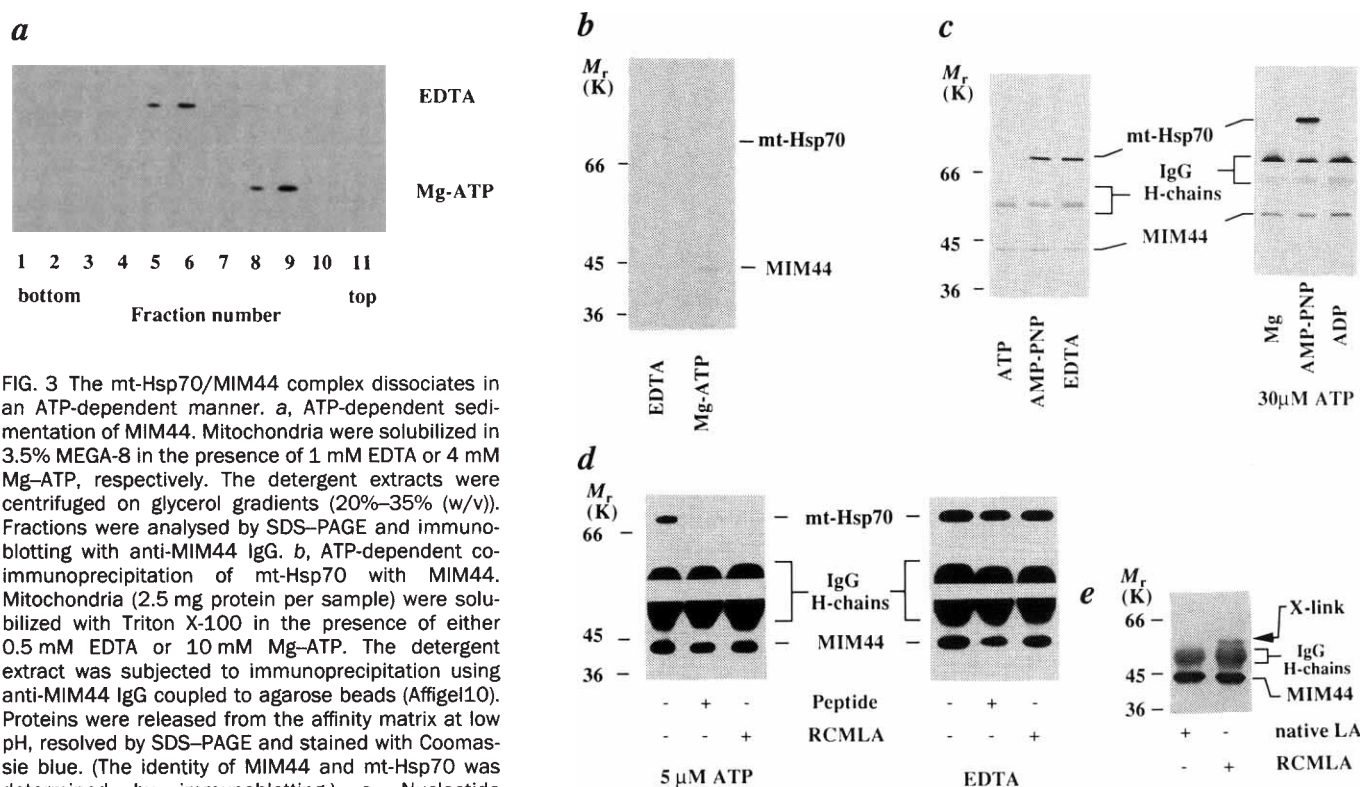


FIG. 3 The mt-Hsp70/MIM44 complex dissociates in an ATP-dependent manner. **a**, ATP-dependent sedimentation of MIM44. Mitochondria were solubilized in 3.5% MEGA-8 in the presence of 1 mM EDTA or 4 mM Mg-ATP, respectively. The detergent extracts were centrifuged on glycerol gradients (20%–35% (w/v)). Fractions were analysed by SDS-PAGE and immunoblotting with anti-MIM44 IgG. **b**, ATP-dependent coimmunoprecipitation of mt-Hsp70 with MIM44. Mitochondria (2.5 mg protein per sample) were solubilized with Triton X-100 in the presence of either 0.5 mM EDTA or 10 mM Mg-ATP. The detergent extract was subjected to immunoprecipitation using anti-MIM44 IgG coupled to agarose beads (Affigel10). Proteins were released from the affinity matrix at low pH, resolved by SDS-PAGE and stained with Coomassie blue. (The identity of MIM44 and mt-Hsp70 was determined by immunoblotting.) **c**, Nucleotide dependence of coimmunoprecipitation of mt-Hsp70 with MIM44. Immunoprecipitations with anti-MIM44 IgG were done in IP-buffer containing 1 mM Mg-ATP, 1 mM Mg-AMP-PNP and 1 mM EDTA, respectively (left panel), or in the presence of 30 μ M Mg-ATP together with 10 mM Mg-acetate, 10 mM Mg-AMP-PNP and 10 mM Mg-ADP, respectively (right panel). Immunoprecipitates from 100 μ g mitochondrial protein were resolved by SDS-PAGE and analysed by immunostaining (using alkaline phosphatase) with antibodies against mt-Hsp70 (upper part) and anti-MIM44 IgG (lower part). **d**, Peptides or unfolded proteins do not disrupt the complex in the presence of EDTA. Mitochondria (100 μ g) were lysed in the presence of 5 μ M ATP and 2 mM Mg-acetate. When indicated the lysis buffer contained RCMLA (60 μ M) or a mixture of two peptides (200 μ M each). Control samples were lysed in the presence of EDTA. **e**, MIM44 binds RCMLA but not native lactalbumin. Mitochondria (100 μ g) were lysed in the

presence of 2 mM Mg-ATP. RCMLA (60 μ M) or native lactalbumin, respectively, were included in the lysis buffer. Samples were subjected to crosslinking with 200 μ M DSS. MIM44-specific crosslinked products were immunoprecipitated and analysed by SDS-PAGE and immunoblotting with MIM44 IgG.

METHODS. **a**, Mitochondrial detergent extracts (1.0 mg protein in 0.4 ml) were loaded on glycerol step gradients (0.57 ml steps of 35%, 32.5%, 30%, 27.5%, 25%, 22.5%, 20% glycerol (w/v) in 3.5% MEGA-8, 30 mM HEPES/KOH, 200 mM K-acetate, 1.0 mM EDTA, 1.0 mM EGTA and 1.0 mM PMSF, pH 7.1) with or without 5 mM Mg-ATP. Mitochondria were lysed in gradient buffer containing 5% (w/v) glycerol. The gradients were centrifuged at 60,000 r.p.m. in a SW 50 Ti rotor (Beckman) for 9 h at 4 °C. **b**, Affinity-purified anti-MIM44 IgG (1 mg) were coupled to Affigel10 beads (0.1 ml bed volume). Before use, the resin was washed with 0.1 M glycine, pH 2.5.

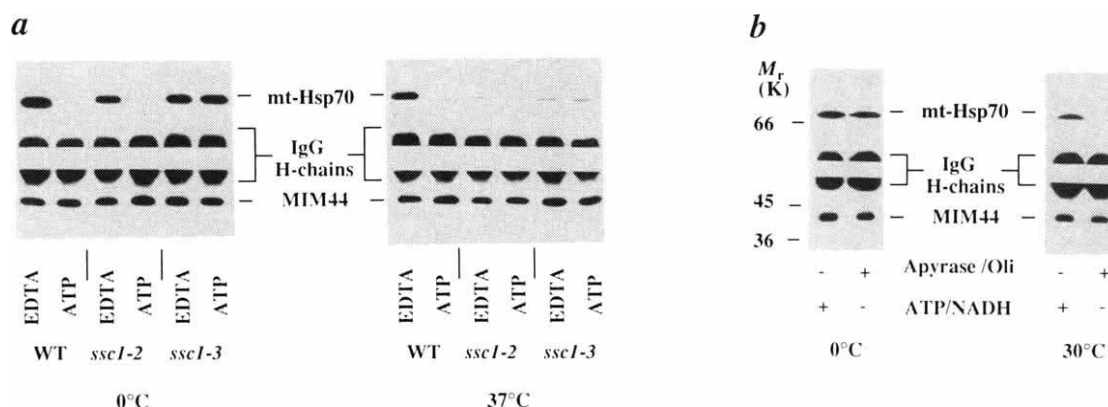


FIG. 4 *a*, Mutant mt-Hsp70s are defective in complex formation with MIM44. Mitochondria harbouring wild-type mt-Hsp70 (WT) or the temperature-sensitive mt-Hsp70s of the *ssc1-2* and *ssc1-3* mutants, respectively, were incubated for 10 min at either 0 °C or 37 °C. After solubilization in the presence of 1 mM EDTA or 10 mM Mg-ATP immunoprecipitation with anti-MIM44 IgG was done. The immunoprecipitates were resolved by SDS-PAGE and analysed by immunoblotting with antibodies against mt-Hsp70 (upper part) and with anti-MIM44 IgG

(lower part). *b*, In intact mitochondria mt-Hsp70 is associated with MIM44 at high but not at low concentration of matrix ATP. Mitochondria (150 µg) were incubated at 30 °C in the presence of apyrase and oligomycin (Oli) to deplete matrix ATP or with added ATP and NADH to maintain high matrix ATP. Control samples were kept at 0 °C. Mitochondria were then lysed and mt-Hsp70 in complex with MIM44 was detected by coimmunoprecipitation and immunostaining.

dissociation of the complex was observed when mitochondria were lysed and briefly incubated in the presence of 5 µM Mg-ATP. When peptides or RCMLA were included, the amount of mt-Hsp70/MIM44 complex was strongly reduced. Native lactalbumine or BSA had no effect.

After ATP-dependent dissociation of the complex, MIM44 could be crosslinked to RCMLA but not to native lactalbumin (Fig. 3e). Without crosslinking RCMLA was not precipitated with anti-MIM44 IgG, suggesting that MIM44 binds unfolded proteins with low affinity.

In summary, the mt-Hsp70/MIM44 complex is maintained in the presence of ATP and disintegrates on ATP hydrolysis. When mitochondria are lysed the complex dissociates in the presence of ATP. It is stabilized when ATP hydrolysis is inhibited by EDTA or competed by AMP-PNP. ADP does not affect ATP-dependent complex dissociation. Unfolded proteins interact with the complex and stimulate its ATP-dependent disintegration. MIM44 in the absence of mt-Hsp70 binds unfolded polypeptides with low affinity.

Complex formation in the mitochondria

Two mutant alleles of the *SSC1* gene that encode temperature-sensitive mt-Hsp70s were described. *Ssc1-2p* carries a mutation in the peptide binding domain, whereas *Ssc1-3p* bears a mutation in the ATPase domain³². At non-permissive temperature, mitochondria harbouring the temperature-sensitive mt-Hsp70s are defective in protein import³².

We analysed the state of association of the mutant mt-Hsp70s with MIM44 (Fig. 4a). At permissive temperature *Ssc1-2p* was found in complex with MIM44 in the presence of EDTA but not of Mg-ATP. By contrast, *Ssc1-3p* was in complex with MIM44 in EDTA and also when Mg-ATP was present in the lysis buffer. Obviously, the nucleotide-dependent dissociation of the MIM44 complex is affected in the *ssc1-3* mutant. When mitochondria were incubated at 37 °C to induce the mutant phenotypes, with both mutants only minor amounts of mt-Hsp70 were recovered in complex with MIM44; this is in contrast to mitochondria from wild type in which the complex was preserved.

To determine the nucleotide dependence of complex formation in intact mitochondria, matrix ATP levels were manipulated²³. Subsequently, mitochondria were lysed in the presence of EDTA, and mt-Hsp70 in complex with MIM44 was monitored by coimmunoprecipitation (Fig. 4b). At high concentration of matrix ATP, mt-Hsp70 was recovered in complex with MIM44, whereas at low matrix ATP concentration only traces were

present. This suggests that the formation of the complex between mt-Hsp70 and MIM44 is inhibited in the absence of ATP. In the mitochondria the complex forms at high ATP levels indicating that ATP hydrolysis by mt-Hsp70 is slow compared to formation of the complex. On ATP depletion the ATP-dependent formation of the complex is reduced shifting the equilibrium towards dissociation. Import of preproteins is blocked in mitochondria from the *ssc1-2* and *ssc1-3* mutant strains³² as well as after depletion of matrix ATP²³. This suggests that formation of the mt-Hsp70/MIM44 complex is crucial for protein import.

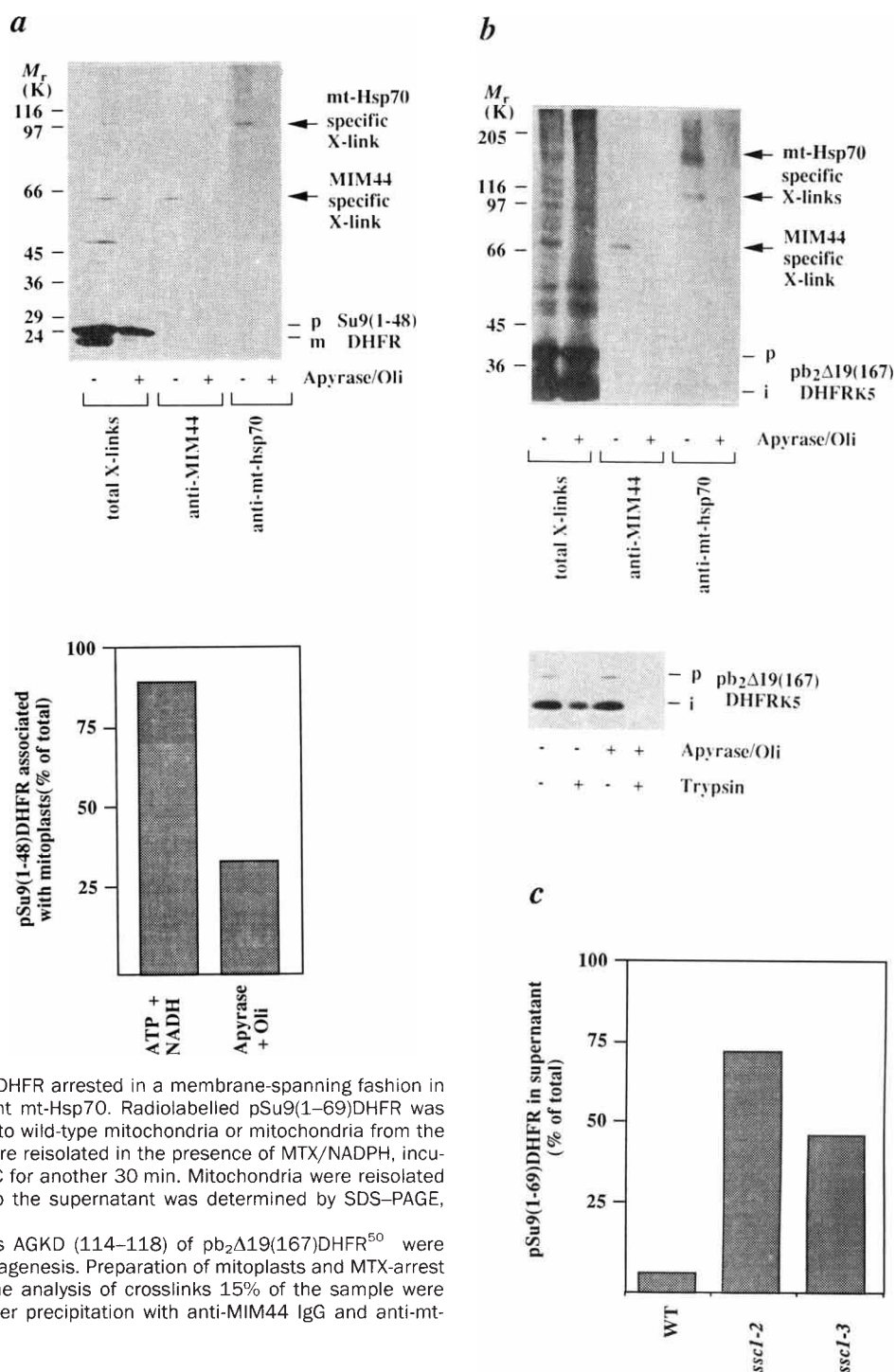
Locking of the translocating chain

To characterize the ATP-dependent interaction of MIM44 and mt-Hsp70 with preproteins we used chemical crosslinking of radiolabelled preproteins arrested at early and late stages of import. The chimaeric preprotein pSu9(1-48)DHFR, consisting of the first 48 amino-acid residues of the ATPase subunit 9 fused to the DHFR, was imported into mitoplasts in the presence of methotrexate (MTX) to stabilize the folded DHFR moiety³³. The translocation intermediate accumulated as the unprocessed precursor form indicating that only the presequence reached the inner face of the inner membrane. Subsequently, ATP levels in the matrix were manipulated and crosslinking was done. In the presence of ATP the spanning intermediate was crosslinked with both MIM44 and mt-Hsp70. On ATP depletion neither of the crosslinks were observed (Fig. 5a, upper panel). In addition, most of the precursor was released into the supernatant (Fig. 5a, lower panel). This agrees with findings³⁴ that arrested preproteins fall out of intact mitochondria if the segment exposed in the matrix is shortened by mitochondrial processing peptidase (MPP) processing to less than ~20 amino-acid residues. These observations demonstrate that, in the presence of the mt-Hsp70/MIM44 complex, a precursor exposing only the presequence into the matrix is held in a spanning fashion. Dissociation of the mt-Hsp70/MIM44 complex on depletion of matrix ATP leads to retrograde movement of the translocating chain and results in release of the preprotein from the mitochondria.

We constructed the fusion protein pb₂Δ19(167)DHFR_{K₅}. This preprotein contains five lysine residues in a position (117 to 121) where they were likely to crosslink efficiently with both MIM44 and mt-Hsp70 when the preprotein was arrested with MTX. After accumulation of pb₂Δ19(167)DHFR_{K₅}, matrix ATP levels were manipulated and crosslinking was done. At high ATP concentration, the arrested translocation intermediate was crosslinked to MIM44 and to mt-Hsp70 but both types of

FIG. 5 Binding of the mt-Hsp70/MIM44 complex to the preprotein in transit prevents its retrograde movement. **a**, Maintenance of the fusion protein pSu9(1-48)DHFR in a membrane-spanning fashion in mitoplasts depends on matrix ATP. Radiolabelled pSu9(1-48)DHFR was imported into mitoplasts for 10 min in the presence of MTX/NADPH. The sample was divided and one half was incubated with NADH/ATP to maintain matrix ATP while the other half was incubated with apyrase and oligomycin to deplete matrix ATP. Aliquots were removed and the fraction of pSu9(1-48)DHFR associated with the mitoplasts was determined. After reisolation of the mitoplasts the rest of the samples were subjected to crosslinking with 200 μ M DSS at 4 °C. Total crosslinks were detected by SDS-PAGE and fluorography. Specific adducts were identified by immunoprecipitation with anti-MIM44 IgG and anti-mt-Hsp70. Upper panel, Analysis of crosslinked species with and without apyrase/oligomycin treatment; lower panel, quantification of the distribution of pSu9(1-48)DHFR between mitoplasts and supernatant. **b**, Crosslinking of the fusion protein pb₂ Δ 19(167)DHFR_{K5} with mt-Hsp70 and MIM44 at high and low concentrations of matrix ATP. Radiolabelled pb₂ Δ 19(167)DHFR_{K5} was imported into isolated mitochondria in the presence of MTX/NADPH. The sample was divided and treated with NADH/ATP and apyrase/oligomycin, respectively. Two aliquots (2%) were removed from each sample and treated with or without trypsin (100 μ g ml⁻¹). The rest of the samples were subjected to crosslinking as in **a**. Upper panel, Analysis of crosslinked species; lower panel, trypsin sensitivity of accumulated translocation intermediate with and without apyrase/oligomycin treatment of mitochondria. Precursor (p) and intermediate sized form (i) of b₂ Δ 19(167)DHFR_{K5} are indicated. **c**, Release into the supernatant of pSu9(1-69)DHFR arrested in a membrane-spanning fashion in mitochondria harbouring wild-type or mutant mt-Hsp70. Radiolabelled pSu9(1-69)DHFR was imported at 25 °C in the presence of MTX into wild-type mitochondria or mitochondria from the *ssc1-2* and *ssc1-3* strains. Mitochondria were reisolated in the presence of MTX/NADPH, incubated for 10 min at 37 °C and then at 25 °C for another 30 min. Mitochondria were reisolated and the fraction of preprotein released into the supernatant was determined by SDS-PAGE, fluorography and densitometry.

METHODS. In pb₂ Δ 19(167)DHFR_{K5} residues AGKD (114-118) of pb₂ Δ 19(167)DHFR⁵⁰ were changed to GGKKKKKD by site-directed mutagenesis. Preparation of mitoplasts and MTX-arrest of preproteins was as described^{33,51}. For the analysis of crosslinks 15% of the sample were applied directly to SDS-PAGE, and 40% after precipitation with anti-MIM44 IgG and anti-mt-Hsp70 IgG, respectively.



crosslinks were strongly reduced on ATP depletion (Fig. 5b). When the protease sensitivity of the translocation intermediate was tested (Fig. 5b, lower panel), the DHFR domain was clipped by trypsin only very slowly at high matrix ATP concentration. This indicates that the folded DHFR domain is tightly apposed to the outer membrane³⁵. In contrast, after depletion of matrix ATP the translocation intermediate became rapidly degraded. This suggests that by lowering matrix ATP, the interaction of the polypeptide chain with the mt-Hsp70/MIM44 complex was relieved and it could slide back in the translocation channel. As a consequence the five lysine residues no longer contact MIM44 or mt-Hsp70 and external protease gains access to the unfolded cytochrome *b*₂ segment.

Finally, we investigated the association of translocation intermediates with mitochondria containing temperature-sensitive

mt-Hsp70s. The fusion protein pSu9(1-69)DHFR was arrested in the presence of MTX in a membrane-spanning fashion. Subsequently the mitochondria were exposed to 37 °C. As shown above (see Fig. 4a) this treatment leads to dissociation of the mt-Hsp70/MIM44 complex in mitochondria from the *ssc1-2* and *ssc1-3* mutant strains but not in wild-type mitochondria. In mitochondria from both mutant strains, pSu9(1-69)DHFR was released into the supernatant whereas only minor amounts of preprotein were released from wild-type mitochondria (Fig. 5c).

In summary all these experiments strongly suggest that formation of the mt-Hsp70/MIM44 complex is required to prevent retrograde movement of preproteins in transit. When mt-Hsp70 function is compromised by lack of ATP or by mutational defects preproteins slide back in the translocation channel. Preproteins which expose only short segments into the matrix fall

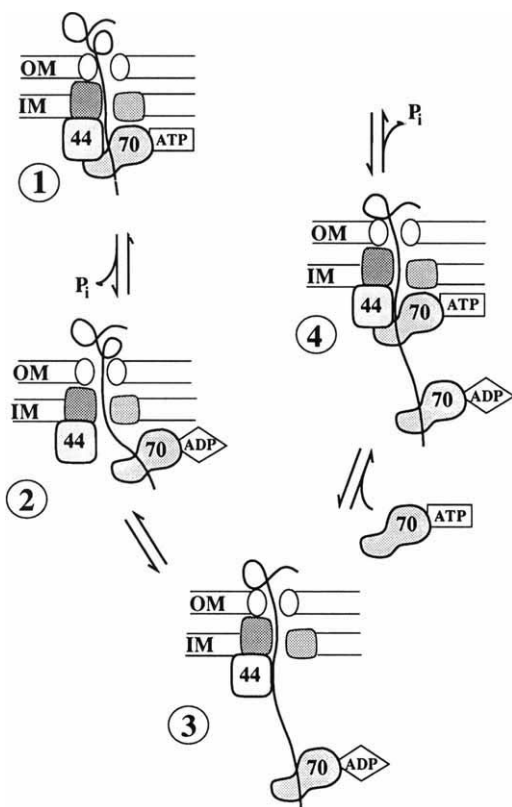


FIG. 6 Model for the cooperation of mt-Hsp70 and MIM44 in protein translocation across the mitochondrial inner membrane. The progression of a translocation intermediate across the outer membrane (OM) and the inner membrane (IM) is schematically outlined. Association-dissociation of mt-Hsp70 (70) and membrane-bound MIM44 (44) is driven by cycles of ATP-hydrolysis. For details see text.

completely out of the mitochondria while longer translocation intermediates slide back but might be held by mt-Hsp70 molecules that are bound to more N-terminal segments of the translocating chain or by partially folded structures in the N-terminal region of the preprotein. Thus, the mt-Hsp70/MIM44 complex has a decisive role in locking the chain in transit to prevent retrograde movement, thereby allowing only forward movement of a preprotein on its way into the matrix.

Discussion

How is the translocation of preproteins across mitochondrial membranes facilitated and how is this process energetically driven? There is good reason to assume that a putative translocation channel, which allows the passage of polypeptides, is coupled to a driving system in which the ATP-dependent binding and release of mt-Hsp70 is an essential process^{17,18}. Mt-Hsp70 appears to exert its function by binding to unfolded polypeptide chains as they appear on the matrix side of the inner membrane, thereby promoting the sliding of the preprotein in an inward direction^{21,24}. In this process mt-Hsp70 conceivably cooperates with components of the translocation apparatus on the inner side of the inner membrane.

We show here that MIM44, a component of the inner membrane¹⁴, represents such a partner for mt-Hsp70. The two components are present in the complex in a roughly 1:1 stoichiometry. The interaction of MIM44 with mt-Hsp70 depends on matrix ATP. MIM44 and mt-Hsp70 can bind to incoming precursors and remain close to the polypeptide chain in transit throughout the import process. Unfolded proteins or peptides accelerate the ATP-dependent dissociation of the mt-Hsp70/MIM44 complex. After complex dissociation, MIM44 binds unfolded polypeptide chains with low affinity.

On the basis of data presented here and of previous studies^{17,18,23,36} we propose a model of how MIM44 and mt-Hsp70 cooperate in the translocation of preproteins (Fig. 6). In this model, MIM44 interacts with the presequence of a preprotein as it appears on the inner face of the inner membrane through the putative translocation channel. At this point mt-Hsp70/ATP interacts with MIM44. In this form, as known from a number of studies on DnaK³⁷ and BiP³⁸, mt-Hsp70 is in a low-affinity state for binding unfolded precursors. We suggest that the ATP on mt-Hsp70 is hydrolysed as the precursor chain switches to the binding site on mt-Hsp70 and stimulates its ATPase activity³⁹. As a consequence, mt-Hsp70 in the ADP form binds the preprotein with higher affinity³⁷, and at the same time the complex of MIM44 with Hsp70 is dissociated. Now the polypeptide chain can slide further through the import channel. The bound mt-Hsp70 would prevent the back-sliding of the polypeptide chain. Then MIM44 rebinds to a more C-terminal segment, and thereby another cycle of mt-Hsp70 binding is initiated. A potential conformational change of MIM44 caused by ATP-dependent interaction with mt-Hsp70 could further facilitate the translocation process. Dissociation of mt-Hsp70 from the preprotein could be triggered by mitochondrial GrpE (MGE), also found in association with the complex. Some mt-Hsp70 may remain bound until folding to the native state occurs or until the polypeptide is transferred to Hsp60²⁴. MGE, in addition to Mdj1p may be involved in these latter reactions comparable to what is known for their bacterial homologues⁴⁰.

In this hypothetical reaction pathway MIM44 has a decisive role in binding the incoming unfolded polypeptide chain and transferring it to mt-Hsp70. The direct regulated interaction of MIM44 with mt-Hsp70 would then provide a molecular machinery which (1) allows mt-Hsp70 to see the unfolded chain as it emerges from the inner membrane and bind it before it has a chance to start folding, (2) confers unidirectionality to the translocation process, and (3) promotes the unfolding of segments of the precursor on the surface of the mitochondria as proposed earlier^{21,22}. Thus, the mt-Hsp70/MIM44 complex could be viewed as a molecular ratchet. Interestingly, a 'brownian ratchet' has been discussed previously on theoretical considerations as a possible principle of membrane transfer of proteins⁴¹.

In the endoplasmic reticulum, BiP, like mt-Hsp70 a member of the Hsp70 family, forms a complex with Sec63p, a DnaJ-related membrane-bound component of the translocation machinery^{42,43}. Although the mechanistic aspects of this interaction are not fully established^{44,45}, it seems conceivable that a complex similar to the one described here could assist in the transmembrane movement of polypeptide chains also in this organelle.

Finally, our results illustrate a new principle of molecular chaperone action. The basic function of a molecular chaperone is to prevent incorrect inter- or intramolecular interactions between transiently exposed surfaces of a protein by binding to them⁴⁶. Here we show that a multifunctional chaperone can be recruited into a complex with a specialized component at a particular location within the cell. Thereby the general ability of the chaperone to bind to exposed unfolded segments of a polypeptide is taken advantage of to facilitate a very specific reaction. Its ability to undergo conformational changes that are regulated by ATP hydrolysis could then be used energetically to drive such a reaction. □

Received 4 July; accepted 22 September 1994.

1. Glick, B. S., Wachter, C. & Schatz, G. *Trends Cell Biol.* **1**, 99–103 (1991).
2. Pfanner, N., Rassow, J., van der Klei, I. & Neupert, W. *Cell* **68**, 999–1002 (1992).
3. Schleyer, M. & Neupert, W. *Cell* **43**, 339–350 (1985).
4. Schwaiger, M., Herzog, V. & Neupert, W. *J. Cell Biol.* **105**, 235–246 (1987).
5. Hwang, S., Jascur, T., Vestweber, D., Pon, L. & Schatz, G. *J. Cell Biol.* **109**, 487–493 (1989).
6. Lill, R., Stuart, R. A., Drygas, M. E., Nargang, F. E. & Neupert, W. *EMBO J.* **11**, 449–456 (1992).
7. Mayer, A., Lill, R. & Neupert, W. *J. Cell Biol.* **121**, 1233–1243 (1993).
8. Kiebler, M., Becker, K., Pfanner, N. & Neupert, W. *J. Membr. Biol.* **135**, 191–207 (1993).

9. Emtage, J. L. T. & Jensen, R. E. *J. Cell Biol.* **122**, 1003–1012 (1993).
10. Dekker, P. J. T. et al. *FEBS Lett.* **331**, 66–70 (1993).
11. Ryan, K. R. & Jensen, R. E. *J. Biol. Chem.* **268**, 23743–23746 (1993).
12. Ryan, K. R., Menold, M. M., Garrett, S. & Jensen, R. E. *Molec. Biol. Cell* **5**, 529–538 (1994).
13. Blom, J. et al. *Molec. cell. Biol.* **13**, 7364–7371 (1993).
14. Maarse, A. C., Blom, J., Grivell, L. A. & Meijer, M. *EMBO J.* **11**, 3619–3628 (1992).
15. Scherer, P. E., Manning-Krieg, U. C., Jenö, P., Schatz, G. & Horst, M. *Proc. natn. Acad. Sci. U.S.A.* **89**, 11930–11934 (1992).
16. Horst, M. et al. *EMBO J.* **12**, 3035–3041 (1993).
17. Kang, P.-J. et al. *Nature* **348**, 137–143 (1990).
18. Ostermann, J. et al. *FEBS Lett.* **277**, 281–284 (1990).
19. Scherer, P. E., Krieg, U. C., Hwang, S. T., Vestweber, D. & Schatz, G. *EMBO J.* **9**, 4315–4322 (1990).
20. Manning-Krieg, U. C., Scherer, P. E. & Schatz, G. *EMBO J.* **10**, 3273–3280 (1991).
21. Neupert, W., Hartl, F.-U., Craig, E. A. & Pfanner, N. *Cell* **63**, 447–450 (1990).
22. Neupert, W. & Pfanner, N. *Phil. Trans. R. Soc. B339*, 355–362 (1993).
23. Stuart, R. A., Gruhler, A., van der Klei, I., Guiard, B., Koll, H. & Neupert, W. *Eur. J. Biochem.* **220**, 9–18 (1994).
24. Stuart, R. A., Cyr, D. M. & Neupert, W. *Trends biochem. Sci.* **19**, 87–92 (1994).
25. Bolliger, L. et al. *EMBO J.* **13**, 1998–2006 (1994).
26. Ikeda, E., Yoshida, S., Mitsuzawa, H., Uno, I. & Toh-e, A. *FEBS Lett.* **339**, 265–268 (1994).
27. Laloraya, S., Gambill, B. D. & Craig, E. A. *Proc. natn. Acad. Sci. U.S.A.* **91**, 6481–6485 (1994).
28. Rowley, N. et al. *Cell* **77**, 249–259 (1994).
29. Rothman, J. E. *Cell* **59**, 591–601 (1989).
30. Langer, T. et al. *Nature* **356**, 683–689 (1992).
31. Palleros, D. R., Welch, W. J. & Fink, A. L. *Proc. natn. Acad. Sci. U.S.A.* **88**, 5719–5723 (1991).
32. Gambill, B. D. et al. *J. Cell Biol.* **123**, 109–117 (1993).
33. Rassow, J. et al. *J. Cell Biol.* **109**, 1421–1428 (1989).
34. Ungermaann, C., Neupert, W. & Cyr, D. M. *Science* (in the press).
35. Schwarz, E., Seytter, T., Guiard, B. & Neupert, W. *EMBO J.* **12**, 2295–2302 (1993).
36. Cyr, M. D., Stuart, R. A. & Neupert, W. *J. Biol. Chem.* **268**, 23751–23754 (1993).
37. Palleros, D. R., Reid, K. L., Shi, L., Welch, W. J. & Fink, A. L. *Nature* **365**, 664–666 (1993).
38. Gething, M.-J., Blond-Eguindi, S., Mori, K. & Sambrook, J. F. in *The Biology of Heat Shock Proteins and Molecular Chaperones* (eds Morimoto, R. I. et al.) 111–135 (Cold Spring Harbor Laboratory Press, New York, 1994).
39. Flynn, G. C., Chappell, T. G. & Rothman, J. E. *Science* **245**, 385–390 (1989).
40. Hartl, F.-U., Hlodan, R. & Langer, T. *Trends biochem. Sci.* **19**, 20–25 (1994).
41. Simon, S. M., Peskin, C. S. & Oster, G. F. *Proc. natn. Acad. Sci. U.S.A.* **89**, 3770–3774 (1992).
42. Brodsky, J. L. & Schekman, R. *J. Cell Biol.* **123**, 1355–1363 (1993).
43. Scidmore, M., Okamura, H. & Rose, M. *Molec. Biol. Cell* **4**, 1145–1159 (1993).
44. Brodsky, J. L. & Schekman, R. in *The Biology of Heat Shock Proteins and Molecular Chaperones* (eds Morimoto, R. I. et al.) 85–109 (Cold Spring Harbor Laboratory Press, New York, 1994).
45. Jungnickel, B., Rapoport, T. A. & Hartmann, E. *FEBS Lett.* **346**, 73–77 (1994).
46. Ellis, R. J. *Phil. Trans. R. Soc. B339*, 257–261 (1993).
47. Daum, G., Böhm, P. C. & Schatz, G. *J. Biol. Chem.* **257**, 13028–13033 (1982).
48. Broach, J. R., Li, Y., Wu, L. C. & Jayaram, M. in *Experimental Manipulation of Gene Expression* (ed. Inouye, M.) 85–109 (Academic, New York, 1983).
49. Hovland, P., Flick, J., Johnston, M. & Sclafani, R. A. *Gene* **83**, 57–64 (1989).
50. Arretz, M., Schneider, H., Guiard, B., Brunner, M. & Neupert, W. *J. Biol. Chem.* **269**, 4959–4967 (1993).
51. Ohba, M. & Schatz, G. *EMBO J.* **6**, 2109–2115 (1987).

ACKNOWLEDGEMENTS. We thank E. A. Craig (Madison, USA) for the kind gift of the *ssc1-2* and *ssc1-3* mutants, M. Meijer (Amsterdam) for a MIM44 cDNA clone and a MIM44 disruption strain, and T. Langer for antiserum against MGE. We thank D. M. Cyr, T. Langer and R. A. Stuart for helpful discussions and M. Moser and C. Zobywalski for excellent technical assistance. This work was supported by grants from the Deutsche Forschungsgemeinschaft Sonderforschungsbereich 184 (Teilprojekt B12), the Human Frontiers of Science Program and the Münchener Medizinische Wochenschrift (to M.B.).

LETTERS TO NATURE

Delayed recombination as a major source of the soft X-ray background

D. Breitschwerdt* & T. Schmutzler*†

* Max-Planck-Institut für Kernphysik, Postfach 103 980, D-69117 Heidelberg, Germany

† Institut für Theoretische Astrophysik, Im Neuenheimer Feld 561, D-69120 Heidelberg, Germany

THE origin of the soft X-ray background radiation has remained mysterious¹ since its discovery², although it is clear from the lack of absorption of the low-energy X-rays that there must be a strong local contribution³. Recent results^{4,5} demonstrate, however, that there are significant more distant contributions, whose origins are also unclear. Here we propose an explanation for both the local and more distant contributions to the soft X-ray background—they seem to arise from the rapid adiabatic expansion of hot gas, driven by the explosions of massive stars. This hot gas cools quickly, ‘freezing in’ highly ionized atomic states. The X-ray emission arises from the delayed recombination of ions and electrons at relatively low temperatures, and is therefore distinct from the more usual line emission excited by collisions with electrons. The X-ray flux is thus relatively insensitive to the local gas kinetic temperature, as the gas is far from ionization equilibrium.

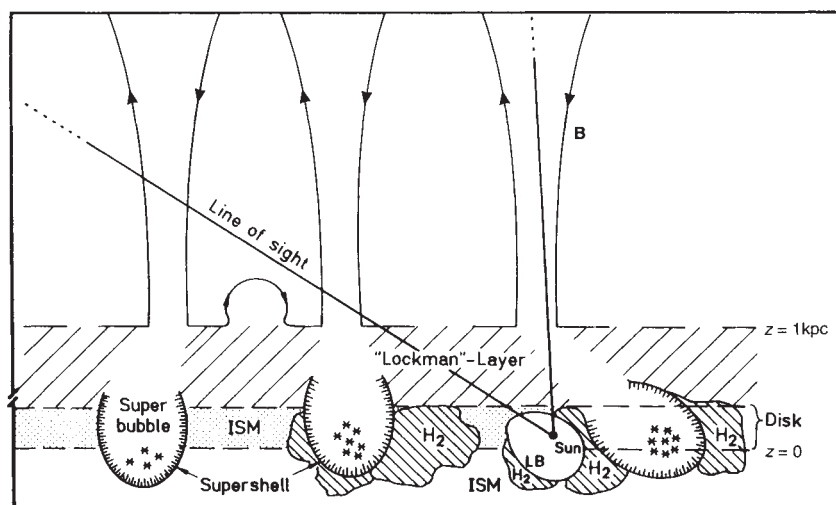
Observations show that associations of young massive stars (OB associations) produce superbubbles which are bound by outward propagating shock waves, compressing and heating swept-up interstellar gas (supershells). Although there is some disagreement whether a supershell breaks out of the Galactic disk⁶, we adopt the widely assumed picture that shock-heated gas is injected into the lower part of the Galactic halo, preferably at places of low surface density (Fig. 1). Presumably these supershells are also sources of cosmic rays, which diffuse out of the disk⁷. The large-scale magnetic field in the Galaxy is mostly parallel to the disk, but becomes inflated by cosmic rays (Parker⁸ instability). Gas and cosmic-ray pressure push the field further out, forming magnetic flux tubes. The escape of cosmic rays from the Galaxy implies a large-scale spatial gradient in the particle distribution, giving rise to resonant generation of magnetohydrodynamic waves⁹. As a consequence of cosmic rays

scattering off these waves, momentum and energy is transferred to the gas¹⁰. We have treated the mutually interacting cosmic rays, gas and self-excited waves as a stationary three-fluid model¹¹ (cosmic-ray-driven Galactic wind model), assuming a flux tube geometry (Fig. 1). Although such a flow is in general time-dependent, a stationary treatment is still appropriate, because our calculations show that the flow time up to the distance where the flow becomes supersonic is of the order of the typical superbubble lifetime.

We fix the boundary conditions (see Fig. 2 legend) at the base of a flux tube, at a vertical distance from the midplane, $z = 1$ kpc. Here the gas temperature and the high ionization states are determined both by the injection of hot superbubble and supershell gas and by type Ia supernova explosions (scale height¹² ~ 700 pc). Because of a continuous input of energy we assume the initial ionization state to be determined by collisional ionization equilibrium at an initial temperature $T_0 = 2.5 \times 10^6$ K. In general the assumption of such an equilibrium, that is a balance between recombination and ionization by collisions with thermal electrons at a given temperature, breaks down at temperatures below 10^6 K, because the radiative cooling timescale becomes much shorter than the relevant recombination timescales. Hence a vast spectrum of timescales in the cooling, ionization and recombination processes determines the ionization states and prevents the gas from reaching ionization equilibrium. A galactic wind flow is characterized by the additional timescale of adiabatic cooling, which can be the shortest one. Time-dependent ionization states affect both the thermal evolution¹³ and the gas dynamics via energy balance. Thus the radiative cooling, depending on the ionization states, determines the dynamics which in turn feeds back to the cooling and hence to the thermal evolution of the gas. Only a self-consistent solution provides the correct information (density, ionization states and temperature) at any point in the flow for calculating emission spectra.

As in models of the solar wind, our results are characterized by a smooth subsonic–supersonic transition in the flow velocity. Thus the mass loss rate of the Galactic wind is fixed by the base velocity of $u_0 \approx 93$ km s⁻¹ at $z = 1$ kpc. The sonic point is reached at $z \approx 18$ kpc, where the flow has been accelerated to $u \approx 194$ km s⁻¹. Although the temperature decreases by two orders of magnitude at a distance of ~ 65 kpc, the high ionization states are essentially retained. Gradual recombination of these ions generates X-ray continuum radiation by the release of

FIG. 1 Schematic edge-on view of the Galactic disk and halo, connected by a layer of neutral hydrogen ("Lockman" layer³⁰). Associations of young massive stars (OB associations) produce superbubbles and supershells. The magnetic field **B** bulges out of the Galactic disk because of Parker instability⁸. Hot gas is injected into the halo by superbubbles breaking out of the gaseous disk, preferably at places of low surface density. Magnetic flux tubes, maintained by the gas and cosmic ray pressure, open up¹¹ significantly at a distance comparable to the Galactic radius (~ 15 kpc). Coupling between cosmic rays and gas by these waves becomes important above $z \sim 1$ kpc. Two lines of sight for a local observer are indicated. (ISM, interstellar medium; LB, Local Bubble.)



energy from high ionization states, predominantly in the puzzling M bands (440–1,100 eV; see Fig. 3 legend). The effect of delayed recombination is clearly demonstrated in the photon emission spectra (Fig. 2a). Note that near 0.75 keV the emission remains at almost the same level in contrast to the decrease at lower energies. This is a striking example of X-ray emission previously not expected from a gas at $\sim 10^4$ K. A theoretical high-resolution spectrum (Fig. 2b) shows conspicuous spikes, which are exclusively due to the recombination continua of the highly ionized species. As a note of caution, we emphasize that neither individual ions nor particular ion ratios provide an

unambiguous indicator of the temperature, because our calculations show that their ratios vary only moderately with temperature.

A spatially integrated spectrum along a flux tube directed towards the Galactic north pole is shown in Fig. 3. At the base of the flux tube the emission in the M bands is dominated by collisionally excited lines, which disappear once the temperature has fallen below $T \sim 10^6$ K, leaving recombination continuum as the only significant contributor. Such behaviour is characteristic of fast adiabatic cooling. We find that gas elements in the flux tube up to a distance of $z \sim 25$ kpc still contribute to the X-ray

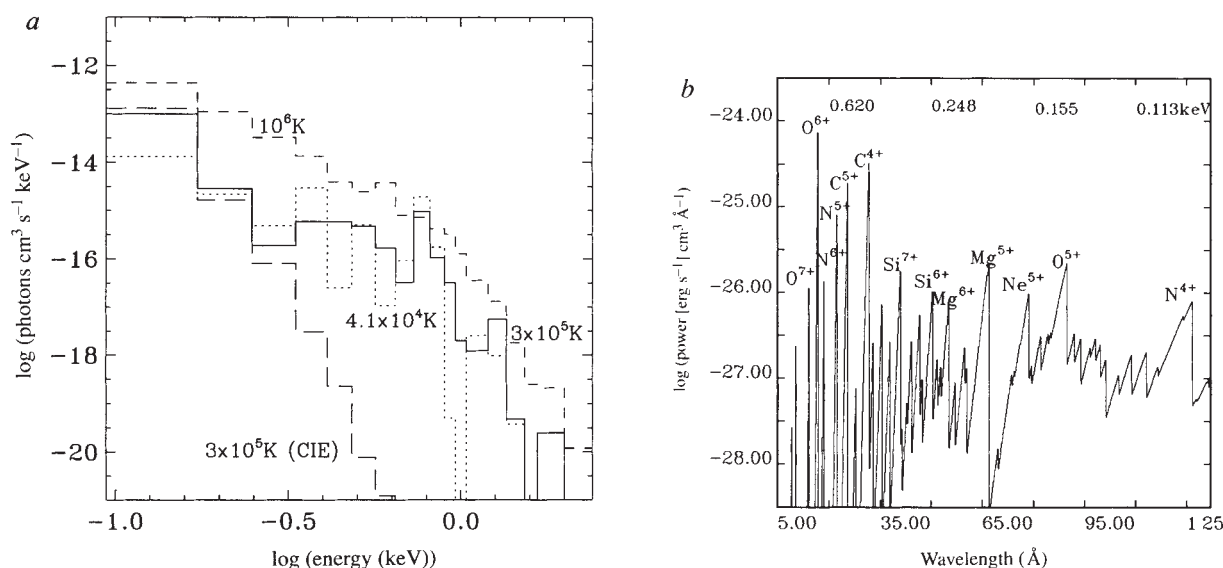


FIG. 2 a, Three spectra of the *in situ* photon emission (normalized by the square of the electron density) in the Galactic wind and, for comparison, a collisional ionization equilibrium (CIE) spectrum. Short-dashed line, the spectrum at $T = 10^6$ K ($z = 10$ kpc); up to $z \approx 25$ kpc, when the gas has cooled to $T = 3 \times 10^5$ K (solid line), it still contributes to the X-ray emission, in contrast to a corresponding CIE spectrum (dashed line). Even at $T = 4.1 \times 10^4$ K ($z \approx 65$ kpc), the spectrum (dotted line) is reminiscent of a 2.5×10^6 K CIE spectrum. The plotted energy range corresponds to the Rosat Position Sensitive Proportional Counter (PSPC) instrument and the binning of channels is arbitrary. The boundary conditions for this model are chosen to match the soft X-ray background. They are fixed at a Galactic radius of $R_0 = 10$ kpc (solar neighbourhood); gas mass density $\rho_0 = 4.2 \times 10^{-27}$ g cm $^{-3}$; initial temperature $T_0 = 2.5 \times 10^6$ K; cosmic ray pressure $P_{c0} = 8.0 \times 10^{-13}$ dyne cm $^{-2}$; regular vertical component of magnetic field $B_0 = 3 \times 10^{-6}$ G; and a random field of $\delta B_0 = 0.1 B_0$. The gas consists of the ten most important elements (H, He, C, N, O, Ne, Mg, Si, S and Fe) with cosmic abundances.

The gravitational potential appropriate for the Milky Way³¹ is a superposition of a bulge, disk and spherically distributed dark-matter halo³² component. The inclusion of a diffuse, averaged photon field (due to Galactic and extragalactic sources³³), appropriate for the solar neighbourhood, has negligible influence on the results. b, High-resolution spectrum (binning in 0.2 Å) of the local emission in the Galactic wind, using the nomenclature of ultraviolet spectroscopy (normalized by the square of the electron number density), at $T = 4.1 \times 10^4$ K corresponding to the same energy range as a. The distinct edges (major contributors are indicated) with exponential decrease to shorter wavelengths are characteristic of recombination continuum. Ionization states represent the non-kinetic or potential part (corresponding to additional degrees of freedom) of the internal energy. Its kinetic part (translational degrees of freedom) is mainly affected by cooling by collisional excitation and consequently operates on different timescales than the non-kinetic part. Therefore a careful inventory of the internal energy¹³ is required.

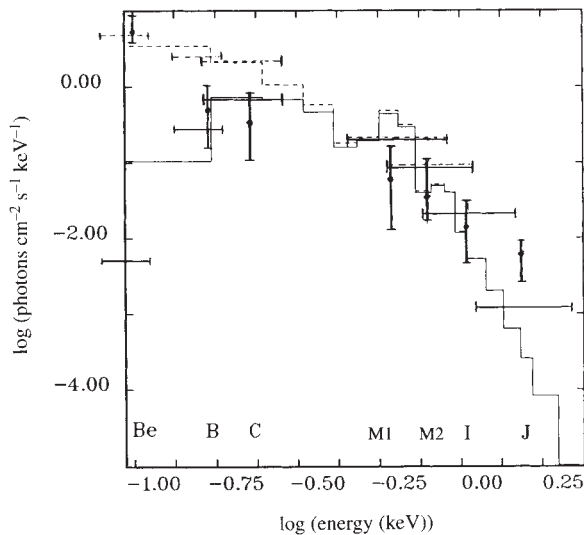


FIG. 3 Photon flux spectra integrated along a flux tube in the Galactic wind model with arbitrary binning (field of view of 2° , appropriate for the Rosat PSPC). The solid line represents the flux reduced by an absorbing, extended H I ('Lockman') layer³⁰ (column density $N_{\text{H}} \approx 1.5 \times 10^{20} \text{ cm}^{-2}$), whereas the dashed line corresponds to a minimum absorption by the local H I cloud²⁵ only, surrounding the Sun (column density $N_{\text{H}} \approx 3 \times 10^{18} \text{ cm}^{-2}$). The horizontal bars show the calculated flux averaged for the following energy bands; Be (77–111 eV), B (130–188 eV), C (160–284 eV), M1 (440–930 eV), M2 (600–1,110 eV), I (770–1,500 eV) and J (1,100–2,200 eV) bands, taking into account the absorption indicated above. The vertical bars give the minimum to maximum range of the measured fluxes taken from the Wisconsin survey¹⁴ in the same energy bands. If other sources contribute (for example, point-like objects), it would be straightforward to relax the boundary conditions of the wind model.

emission, because the emissivity in this energy range remains roughly constant with decreasing temperature (Fig. 2a). Further, our numerical simulations show that the density decreases as $\rho \propto z^{-0.5}$ for $z < 25$ kpc. Thus the observable flux F from the emitting volume, $\Delta V \propto z^2 \Delta z$, varies as $F \propto \rho^2 V / z^2 \approx 1/z$, for constant Δz . In summary, we achieve a quantitative agreement for the B and C bands, consistent with the M bands and even the I band, by a single physical mechanism.

The Wisconsin survey shows that the soft X-ray background radiation is fairly isotropic¹⁴ (in the M bands within an order of magnitude) with respect to Galactic latitude. Measurements from the satellite Rosat suggest that point-like extragalactic sources^{15,16}, presumably active galactic nuclei, contribute roughly half of the M-band emission. However, these sources can only be important for high Galactic latitudes, because absorption by neutral hydrogen increases with decreasing latitude. In our model the isotropy may be explained by (1) a local component, (2) the emission of nearby Galactic superbubbles (at low latitudes) and (3) by integrating the emission along different lines of sight (Fig. 1), that is, intersecting several flux tubes. The patchiness due to isolated flux tubes is smeared out by the limited angular resolution (6.5° full-width at half-maximum) in the Wisconsin survey^{14,17}.

According to observations¹⁸ of H I emission, an average superbubble diameter is ~ 300 pc, a value which we also assume for a typical flux tube. Taking into account the measured spatial variation in the M-band fluxes, we estimate that flux tubes out to a radial distance of 3 kpc dominate the observed emission. The area filling factor of superbubbles¹⁸ is ~ 0.23 , implying ~ 90 flux tubes. This is consistent with the number (13) of known OB associations within 1 kpc² in the solar neighbourhood¹⁹. Even if the whole Galactic disk (radius of 15 kpc) is covered by flux tubes, the total mass loss rate is only ~ 1.8 solar masses per year ($1.8 M_\odot \text{ yr}^{-1}$), comparable to the Galactic star formation rate²⁰.

In order to match the observed isotropy, only half of the superbubbles need to have flux tubes.

Energy-dependent absorption (Fig. 3) suggests a local origin of the ultra-soft X-ray photons (Be to C bands). There is direct observational evidence^{4,5} that the Sun is located in an elongated X-ray-emitting Local Bubble, which extends upwards^{3,31} to ~ 300 pc. Assuming collisional ionization equilibrium, the Be and B' bands (105–188 eV) can be fitted^{3,22} by a temperature of $T \sim 10^6$ K and electron density of $\sim 4.7 \times 10^{-3} \text{ cm}^{-3}$ within an average radius of ~ 100 pc. However, the dispersion measure²³ of the pulsar 0950+08 at a distance of ~ 130 pc yields an average electron density of $n_e \approx 2.3 \times 10^{-2} \text{ cm}^{-3}$. Independently, ultra-violet data on interstellar lines along a different line of sight²⁴ reveal $n_e \approx 2 \times 10^{-2} \text{ cm}^{-3}$ and $T < 5 \times 10^4$ K, implying a mass in the cavity of $\sim 5,000 M_\odot$.

We find that the softest X-ray band ratio, Be to B' band, can be well fitted by the recombination radiation from a fast adiabatically expanding gas (like the Galactic wind); in particular, from a zone characterized by a temperature $T \approx 6 \times 10^4$ K (similar to Fig. 2a). The gas temperature is mainly determined by the degree of adiabatic cooling, whereas the spectrum is the result of the recombination of residual high ionization states. For a line of sight L corresponding to an average radius of the Local Bubble, the calculations show that the required emission measure is $\text{EM} = \int n_e^2 dL \approx 5.8 \times 10^{-2} \text{ cm}^{-6} \text{ pc}$. Using 100 pc this corresponds to an electron density of $n_e \approx 2.4 \times 10^{-2} \text{ cm}^{-3}$. An encouraging aspect of this new class of models, characterized by fast adiabatic cooling, is that it can not only explain the dispersion measure of pulsar 0950+08, but that it also predicts a gas pressure $nT \approx 2,600 \text{ K cm}^{-3}$ (with particle density n) within the Local Bubble, in agreement with both the very local²⁵ and the average interstellar medium. This is a hint that there is no significant expansion of the Local Bubble at present, consistent with velocity measurements of absorption lines (E. B. Jenkins, personal communication).

Therefore we propose that the Local Bubble is the relic of an old superbubble²⁶, created by about ten stars with mass $M \geq 30 M_\odot$ in a dense cloud ($n_{\text{cl}} \sim 10^4 \text{ cm}^{-3}$). It is thus possible to mix in (by turbulence²⁷) a sufficient amount of material from the surrounding shell, resulting in a bubble temperature of $\sim 10^7$ K. After break-out of the cloud, the shell is accelerated by a density gradient in the ambient medium, varying as $\rho_a \propto r^{-\beta}$, where r is the distance from the centre of the bubble. Therefore the radius of the Local Bubble, driven by the most recent supernova explosion ($\sim 4 \times 10^6$ yr ago), is given by $R_{\text{LB}} \propto t^{2/(5-\beta)}$, as long as adiabatic cooling dominates radiative cooling. For $\beta \approx 2$, the density in the cavity decreases as $\rho_{\text{LB}} \propto R_{\text{LB}}^{-3} \propto t^{-2}$. Such an expansion law is similar to the Galactic wind, when the velocity has reached its asymptotic limit and fast adiabatic expansion dominates the cooling. With $R_{\text{LB}} \approx 24$ pc at break-out, the present size of the Local Bubble is reached after $\sim 4 \times 10^6$ yr. Adiabatic cooling reduces the temperature to $T \approx 4 \times 10^5$ K and further, if radiative cooling is included.

The radiative cooling timescale of the Local Bubble at present is sufficiently large ($\sim 5 \times 10^5$ yr) to observe it in its predicted stage of evolution. Recent shadowing experiments⁵ have also revealed local emission in the M bands. Unlike collisional ionization equilibrium and previous nonequilibrium models²⁸ (characterized by ionization delay), both of which conflict with this observation, our model provides a natural explanation of this M-band emission (Fig. 2a). Moreover, we predict the distance of the recently detected binary millisecond pulsar²⁹ (J0437–4715; dispersion measure, $2.65 \text{ cm}^{-3} \text{ pc}$) to be not larger than 110 pc. This could be checked by measurements of parallax. $\square$

Received 5 November 1993; accepted 23 September 1994.

- Cox, D. P. & Reynolds, R. J. *Astr. Astrophys.* **25**, 303–344 (1987).
- Bowyer, C. S., Field, G. B. & Mack, J. E. *Nature* **217**, 32–34 (1968).
- Snowden, S. L., Cox, D. P., McCammon, D. & Sanders, W. T. *Astrophys. J.* **354**, 211–219 (1990).

4. Kerp, J., Herbstmeier, U. & Mebold, U. *Astr. Astrophys.* **268**, L21–L24 (1993).
5. Snowden, S. L., McCammon, D. & Verter, F. *Astrophys. J.* **409**, L21–L24 (1993).
6. Cox, D. P. in *The Interstellar Disk-Halo-Connection* (ed. Bloemen, H.) 143–148 (Kluwer, Dordrecht, 1991).
7. Ginzburg, V. L. & Ptuskin, V. S. *Astrophys. Space Phys. Rev.* **4**, 161–193 (1985).
8. Parker, E. N. *Astrophys. J.* **145**, 811–833 (1966).
9. Lerche, I. *Astrophys. J.* **147**, 689–696 (1967).
10. Wentzel, D. G. A. *Rev. Astr. Astrophys.* **12**, 71–96 (1974).
11. Breitschwerdt, D., McKenzie, J. F. & Völk, H. J. *Astr. Astrophys.* **245**, 79–98 (1991).
12. Tammann, G. A. in *Supernovae* (ed. Schramm, D. N.) 95–116 (Reidel, Dordrecht, 1977).
13. Schmutzler, T. & Tscharnuter, W. M. *Astr. Astrophys.* **273**, 318–330 (1993).
14. McCammon, D., Burrows, D. N., Sanders, W. T. & Kraushaar, W. L. *Astrophys. J.* **269**, 107–135 (1983).
15. Wang, Q. D. & McCray, R. *Astrophys. J.* **409**, L37–L40 (1993).
16. Hasinger, G. *et al.* *Astr. Astrophys.* **275**, 1–15 (1993).
17. McCammon, D. & Sanders, W. T. A. *Rev. Astr. Astrophys.* **28**, 657–688 (1990).
18. Heiles, C. *Astrophys. J.* **354**, 483–491 (1990).
19. Blaauw, A. A. *Rev. Astr. Astrophys.* **2**, 213–237 (1964).
20. Wyse, R. F. G. & Silk, J. *Astrophys. J.* **339**, 700–711 (1989).
21. Paresce, F. *Astr. J.* **89**, 1022–1037 (1984).
22. Juda, M. *et al.* *Astrophys. J.* **367**, 182–185 (1991).
23. Reynolds, R. J. *Astrophys. J.* **348**, 153–160 (1990).
24. Gry, C., York, D. G. & Vidal-Madjar, A. *Astrophys. J.* **296**, 593–599 (1985).
25. Chassefière, E., Bertaux, J. L., Lallement, R., Sandel, B. R. & Broadfoot, L. *Astr. Astrophys.* **199**, 304–311 (1988).
26. McCray, R. & Kafatos, M. *Astrophys. J.* **317**, 190–196 (1987).
27. Kahn, F. D. & Breitschwerdt, D. *Mon. Not. R. astr. Soc.* **242**, 209–214 (1989).
28. Slavin, J. D. & Cox, D. P. *Astrophys. J.* **392**, 131–144 (1992).
29. Johnston, S. *et al.* *Nature* **361**, 613–615 (1993).
30. Lockman, F. J. *Astrophys. J.* **283**, 90–97 (1984).
31. Miyamoto, M. & Nagai, R. *Publs astr. Soc. Japan* **27**, 533–543 (1975).
32. Innanen, K. A. *Astrophys. Space Sci.* **22**, 393–411 (1973).
33. Black, J. H. in *Interstellar Processes Proc. Symp. Interstellar Processes* (eds Hollenbach, D. J. & Thronson, H. A.) 731–742 (Reidel, Dordrecht, 1987).

ACKNOWLEDGEMENTS. We thank P. L. Biermann, F. D. Kahn, W. M. Tscharnuter and H. J. Völk for discussions, and D. P. Cox, for suggestions about the presentation of the manuscript. D.B. and T.S. acknowledge support from the Deutsche Forschungsgemeinschaft.

Path of magnetic flux lines through high- T_c copper oxide superconductors

Zhen Yao, Seokwon Yoon, Hongjie Dai,
Shoushan Fan & Charles M. Lieber*

Division of Applied Sciences and Department of Chemistry,
Harvard University, Cambridge, Massachusetts 02138, USA

A SERIOUS impediment to many potential applications of the high-transition-temperature (high- T_c) copper oxide superconductors is the relative ease with which magnetic flux lines move within these materials, thereby producing finite electrical resistance^{1,2}. To devise methods for rigidly fixing flux lines in these materials, which is necessary to achieve a truly superconducting (zero resistance) state, requires an understanding of their fundamental properties. In clean, conventional type II superconductors, flux lines or vortices can be modelled well as rigid objects that pass straight through a sample. In the high- T_c materials, however, comparatively short coherence lengths, large anisotropies and large accessible thermal energies lead to more complex and fascinating behaviour, giving for example entangled flux lines and two-dimensional pancake vortices^{3–5}. Some detail of the vortex lattice has been resolved previously^{6–13}, although it is not clear how vortices pass through these materials. Here we address this critical issue by simultaneously decorating the positions of flux lines at opposite sides of single-crystal $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_8$ (BSCCO) high- T_c superconductors using the Bitter technique^{14,15}. These new data enable us to quantify the wandering of vortices as they pass through the BSCCO high- T_c materials and address the elasticity of the vortex lattice. This information will be useful for devising effective strategies for pinning flux lines to the crystal lattice.

Conventionally, the Bitter technique involves decorating individual vortices with small magnetic particles as they emerge from one surface of a superconductor^{14,15}. A common weakness of

previous decoration studies is, however, their measurement of only a single two-dimensional slice of the vortex lattice at the sample surface^{7–10}. Because vortices are believed to be flexible in high- T_c materials and because their interaction at the surface differs from that in the bulk, it is difficult to extrapolate convincingly the three-dimensional properties of the vortex lattice from

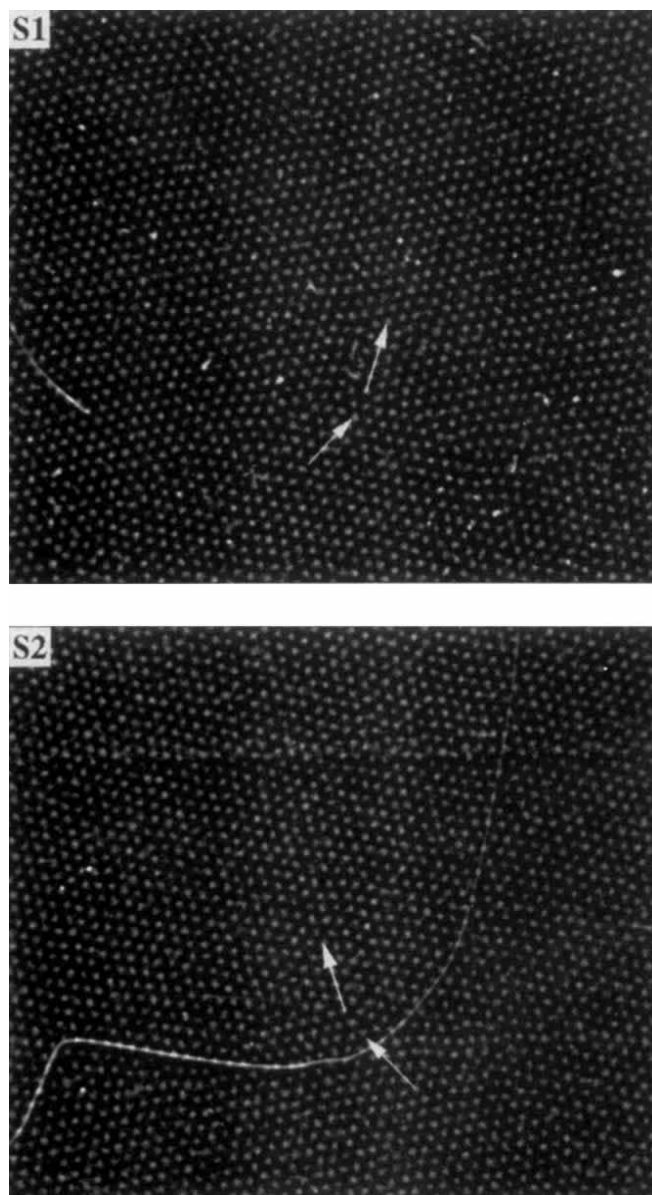


FIG. 1 Scanning electron microscope (SEM) images of side 1 (S1, top) and side 2 (S2, bottom) of a 20- μm -thick BSCCO single crystal decorated with magnetic iron clusters in an applied field of 12 G; S1 and S2 correspond to the same x-y spatial location. The white circular spots in these images correspond to the positions of individual vortices. White arrows in S1 and S2 highlight the different vortex lattice orientations on either side of a grain boundary. Images S1 and S2 are both $65.5 \times 56.1 \mu\text{m}^2$ and the lattice constant is $1.4 \mu\text{m}$. The BSCCO single crystals (typical size, $2 \times 2 \times 0.02 \text{ mm}^3$) were prepared and characterized as described⁹. After cleaving a BSCCO crystal, several hundred ångströms of gold was deposited by sputtering onto each crystal side. Atomic force microscopy showed that the gold films were smooth ($<10 \text{ nm}$ corrugation) and thus do not affect the vortex lattice. The gold-coated samples were mounted with the magnetic field perpendicular to the BSCCO a - b plane, field-cooled to 4.2 K, and then the flux-line positions were decorated with evaporated iron clusters^{7,9}. The arrangement of clusters marking the flux-line positions was imaged at room temperature with a scanning electron microscope. Although decoration is at 4.2 K, the resulting flux-line pattern corresponds to the position of vortices frozen at a higher temperature^{3,9}.

* To whom correspondence should be addressed.

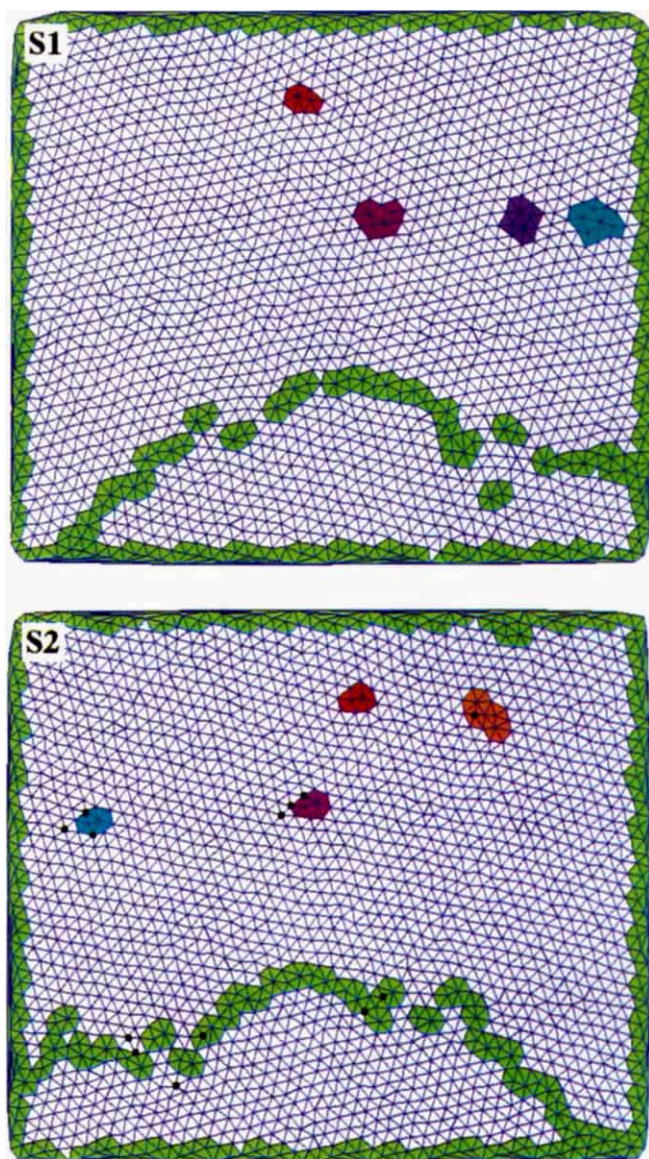


FIG. 2 Delaunay triangulations of the images S1 (top) and S2 (bottom) shown in Fig. 1. In the triangulation, each flux line corresponds to the vertex formed by bonds drawn from the flux line to its n nearest neighbours. Vortices that do not have six nearest neighbours are shaded to highlight their locations in S1 and S2. Apparent extra flux-line positions are indicated by filled black circles on S2.

a single two-dimensional slice^{3,16-18}. Significantly, this limitation of the Bitter technique can be overcome simply by decorating both sample sides simultaneously. Correlation of the vortex lattice structures obtained from both sample sides could then be used to address directly the path of vortices through the sample bulk¹⁷.

A conventional arrangement was used to decorate both sample sides (two-sided decoration), except that the two opposing BSCCO a - b plane surfaces are exposed simultaneously to the source of magnetic particles. Because the contact area of the sample with the thermal sink is greatly reduced when both sample sides are exposed, we found it necessary to reduce the sample heating that occurs during the formation of magnetic clusters. This reduction was achieved by coating both sides of the BSCCO samples with several hundred ångströms of gold. Images of the same x - y location on two sides of a decorated BSCCO sample are shown in Fig. 1. The positions of individual vortices appear as small white spots in these images. We were

able to image the same x - y location to within $\pm 5 \mu\text{m}$ by translating the crystal in the microscope a fixed distance from a reference common to both sides of the sample. This positioning uncertainty is close to but greater than the spacing between flux lines; but by matching localized defects in the vortex lattice (that is, grain boundaries), it is possible to register the two images at the level of individual vortices.

The positions of individual vortices on side 1 (S1) and side 2 (S2) of the sample are well resolved in Fig. 1; two-sided decoration images of similar quality were obtained in independent experiments using the procedures described. Several important facts can be gleaned from these two images. First, the two-dimensional vortex lattice structure observed on either S1 or S2 is similar to that reported previously in single-sided decoration experiments (for similar applied fields)^{8,9}, so we can conclude that the thin gold coating does not affect the observed lattice structure. Second, both S1 and S2 exhibit a grain boundary in the vortex lattice that runs across the lower parts of these images; we highlight the change in lattice direction across the grain boundary using arrows in the images. By placing images S1 and S2 back to back (as they were recorded from the sample), one can see qualitatively that the vortex lattice grain boundaries observed at the two sample sides overlap well and that the lattice orientations in each grain are the same on both sides. Interestingly, we have recently suggested on the basis of single-sided decoration studies of BSCCO that grain boundaries in the vortex lattice should pass through these samples^{9,10}. To our knowledge, our two-sided decoration experiments provide the first proof of this proposal.

We have carried out Delaunay triangulations of the Bitter patterns to obtain greater insight into the three-dimensional vortex lattice structure (Fig. 2). The analysis highlights several significant features. First, the grain boundary already discussed (green) is found to consist of alternating 5-coordinate/7-coordinate lattice sites running along the direction of the grain boundary. The similar topology of this extended defect on both S1 and S2 enables us to register the images to much less than a flux-line lattice constant, a_0 . Second, isolated topological defects are observed on each side of the sample. Three of these defects, which are highlighted in red, magenta and cyan, are common to both sides of the sample and correspond to dislocations in the vortex lattice. The red dislocation is aligned vertically on S1 and S2, whereas the centres of the magenta and cyan dislocations are displaced with respect to one another. In addition, each side exhibits a topological defect (purple, S1; orange, S2) that does

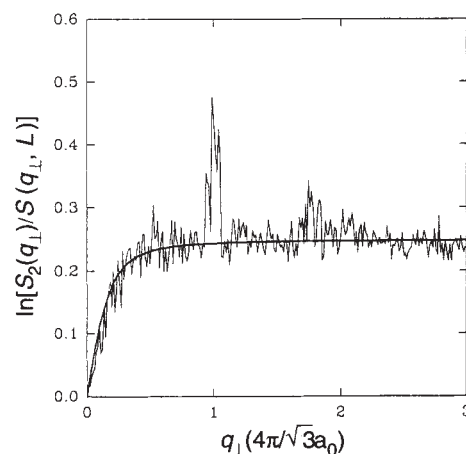


FIG. 3 Plot of $\ln[S_2(q_\perp)/S(q_\perp, L)]$ versus $q_\perp$. The solid curve corresponds to a fit of the data using $L \cdot \sqrt{(c_{11}q_\perp^2/c_{44})/[1 + q_\perp^2\lambda(T)^2]}$, where $c_{11}/c_{44} = 1.5 \times 10^{-4}$ and $\lambda(T) = 1 \mu\text{m}$. Using a value of $\lambda(0) \approx 0.25 \mu\text{m}$ and this value of λ , we estimate that the vortices dropped out of equilibrium at $T = 0.98 T_c$, close to the experimentally observed irreversibility line.

not have a corresponding match on the opposite sample side. The purple defect on S1 is a twisted bond; and probably corresponds to a fluctuation frozen while the sample was cooled before decoration. The orange defect on S2 arises from an interstitial vortex. The observation of these defects implies that at least some vortices wander significantly as they pass through the sample.

This analysis shows that the number of vortices is different on the two sample sides. Extra vortices are highlighted on S2 with solid black dots (Fig. 2). The observation of extra vortices has been confirmed in independent images. These results at first seem remarkable as vortices cannot start or stop within the sample bulk². We believe that the extra vortices can be explained by a small magnetic field inhomogeneity that arises from the non-uniform edges of the BSCCO samples (D. R. Nelson, personal communication). An estimate of this inhomogeneity from the ratio of the number of extra vortex to total vortices yields $\sim 0.6\%$ and is thus quite small. Although the origin of this observation needs to be clarified, our results demonstrate the importance of two-sided decoration in studies of the vortex lattice.

The analysis of topological defects in the vortex lattice at the two sample sides shows that the vortices are flexible and can wander as they pass through the BSCCO samples. We have quantified this wandering by calculating the root-mean-square (r.m.s.) displacement of the flux-line positions $\langle |\mathbf{r}(L) - \mathbf{r}(0)|^2 \rangle^{1/2}$, as they cross the sample from S1 ($z=0$) to S2 ($z=L$), where $\mathbf{r}$ is the flux-line position, L is the sample thickness and the brackets signify an average over all vortices (excluding the extra ones) in the images. Our calculations show that $\langle |\mathbf{r}(L) - \mathbf{r}(0)|^2 \rangle^{1/2} = 0.3 \mu\text{m}$ for the $20\text{-}\mu\text{m}$ thick sample discussed, corresponding to a wandering of $\sim 20\%$ of a_0 . This r.m.s. displacement is comparable to the Lindemann criteria for melting of a solid lattice and again suggests that there is considerable flux-line wandering within these materials.

To determine whether the observed wandering of vortices through these materials affects the fundamental properties of the vortex array, and thus how it might be pinned, we have evaluated a ratio of the vortex elastic constants. The elastic moduli can be determined from the structure function that describes the vortex density fluctuations as a function of the in-plane ($q_{\parallel}$) and out-of-plane ($q_{\perp}$) wavevectors. Here the flux-line positions have been determined at two discrete out-of-plane positions ($z=0$ and $z=L$, where L is the sample thickness), and in this case the structure function can be written $S(q_{\perp}, L) = \langle \rho_{q_{\perp}}(L) \cdot \rho_{q_{\perp}}^*(0) \rangle$, where $\rho_{q_{\perp}}$ is the Fourier transform of the vortex density in a given plane. The structure function has been evaluated previously^{3,18,19}, showing that

$$S(q_{\perp}, L) = S_2(q_{\perp}) \exp(-\varepsilon(q_{\perp})L/k_B T) \quad (1)$$

where $S_2(q_{\perp})$ is computed for a constant z slice ($z=0$ or $z=L$). By treating explicitly the repulsive interaction between vortices

$$\frac{\varepsilon(q_{\perp})}{k_B T} \approx \sqrt{\frac{c_{11}q_{\perp}^2/c_{44}}{1 + q_{\perp}^2 \lambda(T)^2}} \quad (2)$$

where c_{11} and c_{44} are the $\mathbf{q} \rightarrow 0$ bulk and tilt moduli and $\lambda(T)$ is the temperature-dependent penetration depth (for our magnetic field and sample size, $c_{11}(\mathbf{q}) \approx c_{11}/(1 + q_{\perp}^2 \lambda^2)$ and $c_{44}(\mathbf{q}) \approx \text{constant}$ ²⁰). Physically, equations (1) and (2) imply that the decay in flux-line correlations through the sample is governed by the bulk and tilt moduli of the vortices. This model was originally derived for flux liquids, but equation (2) has the same form for a thermally excited crystalline lattice away from the reciprocal lattice vectors, provided $c_{66} \ll c_{11}(\mathbf{q})$ (ref. 19). Hence our analysis is not limited to a particular state of the vortices.

As $S(q_{\perp}, L)$ and $S_2(q_{\perp})$ can be calculated from our two-sided decoration results, it is possible to evaluate readily c_{11}/c_{44} and $\lambda(T)$ from equation (2) (Fig. 3). Several important conclusions can be drawn from these results. First, we find that $\varepsilon(q_{\perp})$ goes linearly to zero in the limit $q_{\perp} \rightarrow 0$. This result shows that we

are dealing with line-like objects as opposed to two-dimensional pancake vortices¹⁷. Second, we can fit the experimental data very well over the range of experimentally accessible wavevectors, and thus reliably determine that $c_{11}/c_{44} = 1.5 \times 10^{-4}$ and $\lambda(T) = 1 \mu\text{m}$. Significantly, the small value of c_{11}/c_{44} determined in these studies suggests a strong downward renormalization of the bulk modulus from its zero temperature value; that is, the vortex array is much more easily compressed than expected for rigid parallel lines²⁰. This strong renormalization cannot be explained by simple crystal defects such as dislocations in a crystalline lattice, but may be due to thermal wandering of vortices from straight trajectories^{3,19}. Although more theoretical and experimental work will be needed to understand this fundamental renormalization of c_{11} , we believe that this new information will be useful in devising strategies for effectively pinning vortices and thereby enhancing critical current in the high- T_c materials. $\square$

Received 15 July; accepted 21 September 1994.

1. Bishop, D. J., Gammel, P. L., Huse, D. A. & Murray, C. A. *Science* **255**, 165–172 (1992).
2. Huse, D. A., Fisher, M. P. A. & Fisher, D. S. *Nature* **358**, 553–559 (1992).
3. Nelson, D. R. & Seung, H. S. *Phys. Rev. B* **39**, 9153–9174 (1989).
4. Brandt, E. H. J. *Supercond.* **6**, 201–217 (1993).
5. Clem, J. R. *Phys. Rev. B* **43**, 7837–7846 (1991).
6. Dolan, G. J., Chandrasekhar, G. V., Dinger, T. R., Feild, C. & Holtzberg, F. *Phys. Rev. Lett.* **62**, 827–830 (1989).
7. Griër, D. G. et al. *Phys. Rev. Lett.* **66**, 2270–2273 (1991).
8. Bolle, C. A. et al. *Phys. Rev. Lett.* **66**, 112–115 (1991).
9. Dai, H., Liu, J. & Lieber, C. M. *Phys. Rev. Lett.* **72**, 748–751 (1994).
10. Yoon, S., Dai, H., Liu, J. & Lieber, C. M. *Science* **265**, 215–218 (1994).
11. Cubitt, R. et al. *Nature* **365**, 407–411 (1993).
12. Yethiraj, M. et al. *Phys. Rev. Lett.* **70**, 857–860 (1993).
13. Harada, K. et al. *Phys. Rev. Lett.* **71**, 3371–3374 (1993).
14. Essman, U. & Trüble, H. *Phys. Status Solidi* **18**, 813–828 (1966).
15. Huebner, R. P. *Magnetic Flux Structures in Superconductors* (Springer, Berlin, 1979).
16. Huse, D. A. *Phys. Rev. B* **46**, 8621–8623 (1992).
17. Nelson, D. R. in *Phenomenology and Applications of High-Temperature Superconductors* (eds Bedell, K. S., Inui, M., Meltzer, D., Schrieffer, J. R. & Doniach, S.) 187–242 (Addison-Wesley, New York, 1992).
18. Marchetti, M. C. & Nelson, D. R. *Phys. Rev. B* **47**, 12214–12223 (1993).
19. Nelson, D. R. & Le Doussal, P. *Phys. Rev. B* **42**, 10113–10129 (1990).
20. Fisher, D. S. in *Phenomenology and Applications of High-Temperature Superconductors* (eds Bedell, K. S., Inui, M., Meltzer, D., Schrieffer, J. R. & Doniach, S.) 287–327 (Addison-Wesley, New York, 1992).

ACKNOWLEDGEMENTS. We thank D. R. Nelson for discussion. C.M.L. acknowledges support of this work by the NSF Division of Materials Research and the Harvard Materials Research Laboratory.

Eemian climate fluctuations observed in a European pollen record

Michael H. Field*, Brian Huntley† & Helmut Müller‡

* Sub-Department of Quaternary Research, University of Cambridge, The Botany School, Downing Street, Cambridge CB2 3EA, UK

† Environmental Research Centre, University of Durham, Department of Biological Sciences, South Road, Durham DH1 3LE, UK

‡ Bundesanstalt für Geowissenschaften und Rohstoffe, Zi.X32, Postfach 51 01 53, D 30631, Hannover 51, Germany

RECENT ice-core data from Greenland^{1,2} suggest that the climate during the last interglacial period (the Eemian) was more unstable than that of the Holocene (about 10,000 years ago to the present), being characterized in particular by a series of cold episodes each lasting about 70 to 750 years. Subsequent analysis of a second Greenland ice core^{3,4}, however, failed to corroborate the details of these Eemian climate fluctuations, a result that may be attributable to the effects of ice flow⁴. To resolve this discrepancy, it is imperative to seek alternative sources of information about the Eemian climate. Here we present climate reconstructions from pollen data

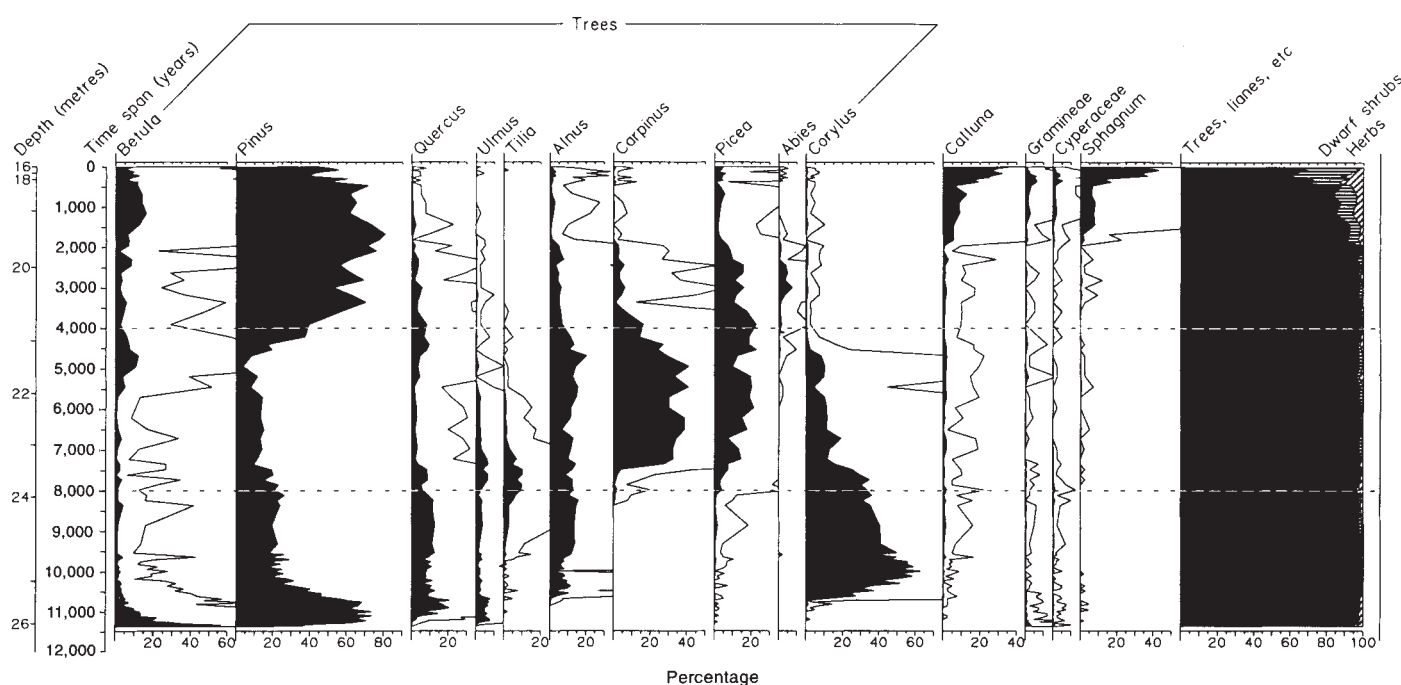


FIG. 1 Summary pollen diagram for Bispingen, northwest Germany (53° 5' N, 9° 59' E, ~100 m above sea level). Relative abundance values of the more important pollen and spore taxa plotted with respect

to estimated time span. (Unfilled area under curve shows $\times 10$ scale exaggeration; modified after Müller⁸).

from the annually laminated Eemian lake-sediment record at Bispingen⁵ and from the Eemian and Holocene peat records at La Grande Pile⁶. The former record indicates that an initially warm period of 2,900 yr was followed by cooling and a series of colder episodes, one of which had winter temperatures comparable to those at the end of the preceding cold stage. The latter records show greater climate instability during the Eemian than the Holocene. These results are in broad agreement with those from the GRIP ice core, but contrast both with the GISP2 core^{3,4} and with recent high-resolution marine records from the North Atlantic^{7,8}.

Last interglacial (Eemian) deposits occur widely in Europe, being identified by their stratigraphic position and pollen stratigraphy⁹. The Eemian correlates with the Ipswichian (Britain), Sangamon (North America) and Mikulino (Russia), and with marine oxygen isotope stage 5e (ref. 9). A pollen diagram (Fig. 1) from lake sediments at Bispingen, northwest Germany, shows features typical of the Eemian interglacial. Although the initial rise in tree pollen values is missing, the lowermost samples having abundant *Betula* pollen, comparison with La Grande Pile⁶ suggests only a short period passed before organic sedimentation began in the lake. Decreasing tree pollen abundance and increases in heathland taxa mark the end of the interglacial.

The occurrence of annual laminations⁵ at Bispingen enables a timescale to be assigned (Table 1) allowing estimation of rates of climate change and duration of climate events. Although countable laminations are not present throughout (Table 1), uniform sediment composition allows extrapolation of accumulation rates⁵. The Eemian ends 1,500 yr below the uppermost pollen sample. The majority of the stage is present and spans 9,700 yr, a figure in close agreement with other estimates¹. Using this timescale, we calculate rates-of-change of Eemian pollen spectra composition (Fig. 2), allowing comparison with analyses of Holocene data from North America¹⁰ and Europe^{11,12}. Rate-of-change maxima indicate times of rapid change in pollen spectra and, by implication, in past vegetation. Highest rates occur at the transition from boreal to thermophilous deciduous tree pollen dominance. A similar peak is observed at the transition from the preceding glacial to the Holocene^{10,12}. Holocene data

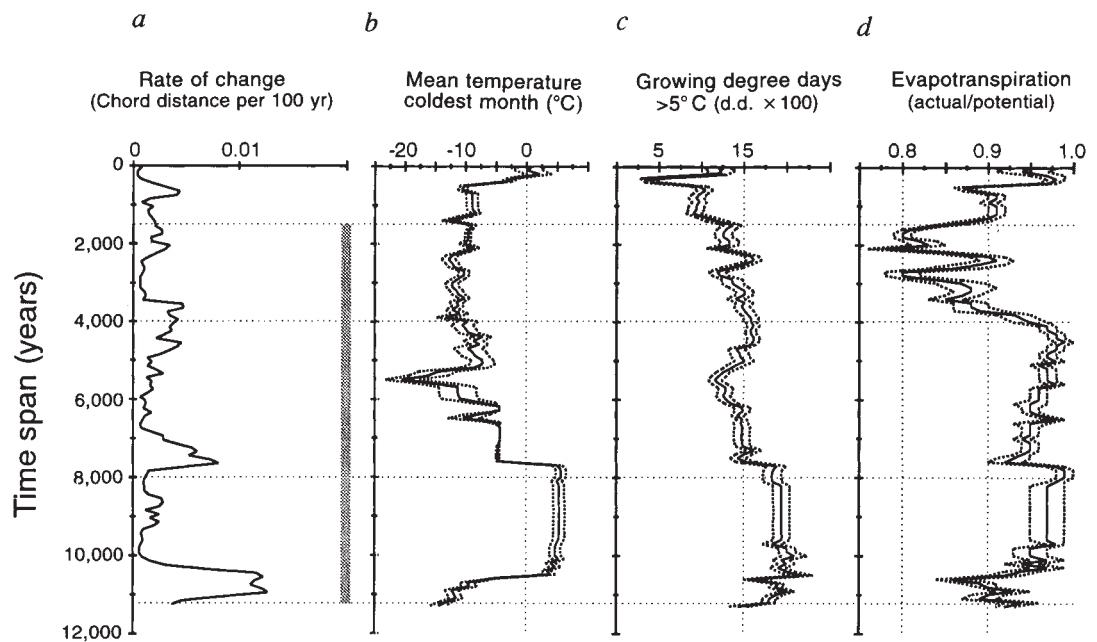
subsequently exhibit periods of stability, with intervening peaks related to principal pollen-stratigraphic changes^{10,12}. The present record exhibits peaks of similar magnitude to those in Holocene data, but only short intervals of values as low as those during most of the Holocene¹².

TABLE 1 The Bispingen timescale

Depth (m)	Cumulative 'age' (yr)	Time interval (yr)	Basis of estimate (after Müller ⁵)
16.65	0	—	Topmost sample in pollen stratigraphy
19.10	400	400	Sediment accumulation rate of 6 mm yr ⁻¹ estimated from count of 100 annual laminations in 60-cm interval of sediment within this lithological unit
20.55	2,400	2,000	Constant sediment accumulation rate assumed, based on the constant proportion of diatomite in these sediments compared to underlying sediments
21.65	4,400	2,000	As above
24.00	8,400	4,000	As above
24.98	9,500	1,100	Extrapolation of sediment accumulation rate determined from counts made on short sequences of very fine annual laminations
25.57	10,200	700	Counted annual laminations
25.93	10,650	450	As above
26.40	11,100	450	As above
26.61	11,300	200	As above
26.68	11,400	100	Estimate from discontinuous count of annual laminations

The relative 'age' of each pollen sample, expressed in years relative to the topmost sample, was obtained by linear interpolation between the 11 control points tabulated above. These are based on counted annual laminations and derived estimates of sediment accumulation rates.

FIG. 2 a–d, Eemian rate-of-change and palaeoclimate reconstruction for Bispingen plotted with respect to estimated time span. a, Rate-of-change of composition of pollen spectra. Pollen values were linearly interpolated at regular 100-yr intervals giving 114 pollen spectra. The chord distance was calculated between consecutive samples to provide an assessment of rate of change^{10,11}. Aquatic and lower plant pollen/spore taxa that primarily reflect local ecological changes within the lake were excluded when making these calculations; all identified terrestrial pollen/spore taxa were included. b, Mean



temperature of the coldest month (MTCO). Climate reconstructions were made for the interpolated pollen spectra using pollen–climate response surfaces^{15–17} for 28 pollen taxa²⁰. The response surfaces were fitted to >7,000 pollen surface samples from western Eurasia and North America. Values were reconstructed for the three 'bioclimate' variables that most strongly influence terrestrial vegetation in temperate latitudes²⁸, values for which were interpolated for each surface sample locality^{17,20}. Analogues selected as the basis for climate reconstruction were constrained to represent the same 'biome'; biome assignment used the procedure of Guiot *et al.*¹⁹. c, Annual temperature sum above

a threshold of 5 °C (GDD5) (d.d., degree days). d, Actual to potential evapotranspiration ratio (AET/PET). For a–d, reconstructed values are based on the 10 closest analogues to each fossil pollen spectrum. Solid lines in b, c and d represent distance-weighted means of the values for these 10 closest analogues, dotted lines show lower and upper 95% confidence intervals derived from the standard errors of these means. Upper and lower guidelines indicate the beginning and end of what we regard as the interglacial, as indicated by the bar adjacent to the rate-of-change curve. Interpolated present-day values for the three bioclimate variables are –0.3 °C, 1,681 degree days and 1.0.

As in previous studies^{13,14}, our quantitative palaeoenvironmental reconstructions are based on surface sample data. We use climate response surfaces^{15–17} and measure degree of analogy between fossil and modern pollen spectra using chord distance¹⁸. We constrain potential analogues to those representing the same biome as the fossil spectrum¹⁹. Closest analogues to Eemian and Holocene spectra show similar chord distances; fewer 'no-analogue' problems are presented than by late-glacial pollen spectra²⁰. Fitting pollen–climate response surfaces^{15–17} reveals that most taxa respond nonlinearly to individual climate variables and generally show interacting responses to several variables²¹.

The Bispingen palaeoclimate reconstruction begins with 100 yr of marked temperature seasonality and low mean temperature of the coldest month (MTCO) analogous to climate today in central European Russia. A rapid increase in MTCO from ~–13 to ~5 °C in ~700 yr follows, beginning 11,200 yr below the top of the profile and marking the beginning of the interglacial; this increase is comparable in rate and magnitude to a reconstructed winter temperature increase in Britain from ~–17 °C at 10,500 yr before present (BP) to 0–1 °C at 9,800 yr BP (refs 20, 22). The next 2,900 yr of the Eemian had reduced seasonality and greater moisture availability; comparable conditions occur today in northwestern France. MTCO and annual temperature sum above 5 °C (denoted GDD5) subsequently decreased at a relative age of 7,600 yr, leading to conditions similar to southern Sweden today. Such conditions persisted for 1,500 yr, although interrupted by a cooling episode of 200–300 yr centred at 6,500 yr. A major oscillation in winter temperatures occurred during the millennium after 6,100 yr. At its extreme the climate perhaps resembled that today in central Siberia. This oscillation corresponds to the decline or disappearance of several thermophilous taxa (*Corylus*, *Fraxinus*, *Taxus*,

Tilia, *Ulmus*) with geographical limits related to winter temperatures; *Corylus* pollen values are at their lowest since increasing at ~10,500 yr whereas those for cold-tolerant *Betula* increase. Thus, although the analogues for this time are the poorest of those used, evidence for winter temperature decrease is clear. Between 6,100 and 5,100 yr temperature seasonality increased relative to the interval preceding the oscillation and conditions became steadily moister, culminating in a climate like that of northern Belarus today. Increasingly dry conditions between 5,100 and 3,700 yr lead to a climate like that today in southwestern European Russia. A period (1,500 yr) of rapidly fluctuating moisture conditions followed before a trend of decreasing temperature seasonality and increasing moisture availability at 2,200 yr. This trend continued for 700 yr leading to conditions like those today in coastal northern Norway. The last 1,500 yr of the profile show rapid changes in all three climate variables. The topmost sample indicates relatively moist conditions with mild winters comparable to southeast Sweden today. Recurrence of pollen of thermophilous trees in these sediments, however, may result from re-working of earlier deposits. This might reflect water level change in the basin or result from permafrost activity; the climate of the preceding interval is consistent with establishment at least of discontinuous permafrost. The climate fluctuations reconstructed from these uppermost samples thus may be spurious, although high *Calluna* pollen abundance is inconsistent with very low temperatures and unlikely to result from re-working.

We applied the same method to reconstruct climate from the pollen record at La Grande Pile⁶. This reconstruction shows rapid increases in MTCO of ~28 °C and ~20 °C at the onset of the Eemian and Holocene, respectively, (Fig. 3) agreeing broadly with earlier reconstructions from this site¹⁵. Similar magnitudes of mean annual temperature increase at the onset of the Eemian

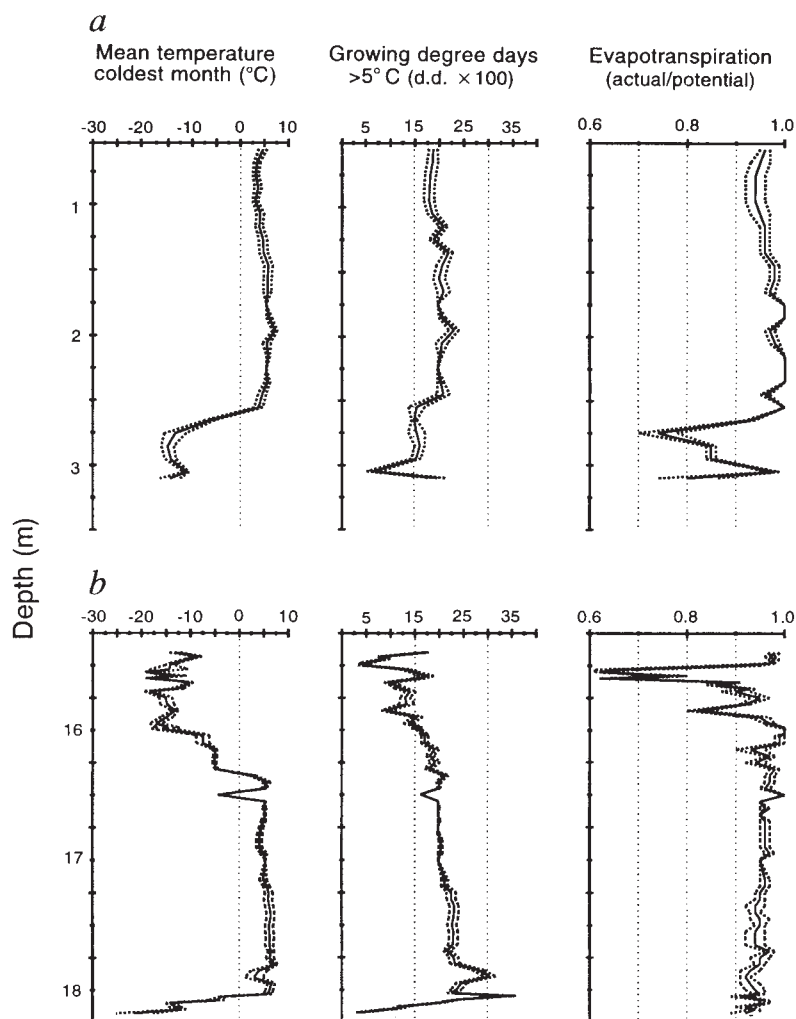


FIG. 3 Holocene (a) and Eemian (b) palaeoclimate reconstructions for La Grande Pile, France, plotted with respect to depth. Reconstructions were made using the same technique as for the Bispingen data, except that the original pollen spectra were used rather than spectra interpolated for regular intervals of time/depth. a, MTCO, GDD5 and AET/PET for the Holocene. b, MTCO, GDD5 and AET/PET for the Eemian. (Interpolated present-day values for the three bioclimate variables are 0.6 °C, 1,906 degree days and 1.0).

and Holocene have been reconstructed from Les Echets¹⁴. These results support those from Bispingen, allowing us to conclude that climate changes in Europe during Saalian and Weichselian glacial terminations were similar both in magnitude and rate. In addition, MTCO values reconstructed for the Saalian and Weichselian glacial maxima at La Grande Pile are similar (~ -35 °C) and indicate additional less rapid warming of 10–15 °C before the phase of rapid warming. Although no timescale can be attributed at La Grande Pile, the climate reconstruction indicates similar overall seasonality changes during the Eemian to those at Bispingen, albeit with the early less-seasonal phase occupying a large part of the depth spanned by the interglacial, perhaps as a result of a change in sedimentation rate.

Thus the palaeoclimate reconstructions reveal similar rates and magnitudes of climate change at the transitions to the Eemian and Holocene. The climatic sequences within the interglacials differ, however, the Eemian showing an early phase of low seasonality followed by increasing temperature seasonality whereas the early Holocene, particularly in northern Europe, is relatively seasonal with minimum seasonality during the late Holocene²³. These differences between the sequences of climate change are consistent with differences in seasonal insolation patterns between the two warm stages⁹. Thus the distinctive pollen-stratigraphic 'signatures' of each Pleistocene warm stage, including the Holocene, probably reflect differences in climate that are in turn due to the continuously varying insolation regime^{9,24}.

The relative stability of Holocene climate at La Grande Pile also contrasts with that of the Eemian at the same site. Although the latter reconstruction does not display the extreme episodes seen at Bispingen, it does show markedly decreased MTCO asso-

ciated with greater overall instability following an initial phase of relative stability with high MTCO and relatively high GDD5 values. Unless Bispingen is exceptional in some way, lack of evidence for the more extreme climatic episodes at La Grande Pile might reflect either insufficient temporal resolution of this record or, more likely, that its more southerly and westerly position rendered it less susceptible to whatever change in atmospheric circulation is reflected by the episodes recorded at Bispingen. This suggests that hypotheses to explain Eemian climate fluctuations should focus on mechanisms giving marked winter cooling in northern Europe; one plausible scenario invokes Norwegian Sea surface cooling and sea-ice expansion with consequent intensification of the Fennoscandian winter anticyclone.

Detailed comparisons with Greenland ice-core data are rendered difficult both by lack of a direct seasonal temperature reconstruction from Greenland and by the tentative nature of the timescales estimated for those cores^{1,4}. However, two points of similarity do emerge. First, both continental pollen records and the GRIP ice-core data^{1,2} indicate less stable climate during the Eemian stage than during the Holocene. Second, at Bispingen, the largest of the climatic oscillations is of such a magnitude that, at least for MTCO, values drop to levels comparable with those preceding the Saalian glacial termination. This apparently parallels inferred cooling to 'mid-glacial climate conditions' in Greenland at several points during the Eemian¹. Reconstructions from continental pollen data, however, more readily reveal aspects of seasonality. The Bispingen results indicate most marked fluctuations in winter temperatures; the modest GDD5 decrease implies that summer temperatures remained

relatively warm during even the extreme oscillation. This is not consistent with the inference of 'mid-glacial climate conditions' from ice-core data¹. Instead, the Bispingen data focus attention on a mechanism that would lead to marked winter cooling without substantial impact upon summer conditions. Palaeoclimate reconstructions for the British Isles during the Weichselian late-glacial^{20,22} indicate much larger winter than summer temperature changes. These changes are generally regarded as the consequence of migration of the Polar Front in the North Atlantic²⁵ and have been linked to changes in thermohaline circulation of the oceans²⁶. Thus it is tempting to speculate that Eemian climate fluctuations resulted also from this mechanism. Recent results from an idealized ocean model²⁷ support this suggestion. The relatively high July insolation and consequent intensified hydrological cycle of the early Eemian may have induced instability in North Atlantic circulation. This could lead to episodes with cold winters in northern Europe coupled to relatively warm summers maintained by the direct effect of insolation. Our reconstructions are consistent with such a pattern. □

Received 18 August, accepted 21 September 1994.

1. GRIP Members *Nature* **364**, 203–207 (1993).
2. Dansgaard, W. *et al.* *Nature* **364**, 218–220 (1993).
3. Taylor, K. C. *et al.* *Nature* **366**, 549–552 (1993).

4. Grootes, P. M., Stuiver, M., White, J. W. C., Johnsen, S. & Jouzel, J. *Nature* **366**, 552–554 (1993).
5. Müller, H. *Geol. Jb.* **A21**, 149–169 (1974).
6. Woillard, G. M. *Quat. Res.* **9**, 1–21 (1978).
7. Keigwin, L. D., Curry, W. B., Lehman, S. J. & Johnsen, S. *Nature* **371**, 323–326 (1994).
8. McManus, J. F. *et al.* *Nature* **371**, 326–329 (1994).
9. Watts, W. A. in *Vegetation History* (eds Huntley, B. & Webb, T.) 155–192 (Kluwer Academic, Dordrecht, 1988).
10. Jacobson, G. L., Webb, T. & Grimm, E. C. in *North America and Adjacent Oceans during the Last Deglaciation* (eds Ruddiman, W. F. & Wright, H. E.) 277–288 (Geological Soc. of America, Boulder, Colorado, 1987).
11. Huntley, B. *Clim. Dyn.* **6**, 185–191 (1992).
12. Kelly, M. G. & Huntley, B. *J. Quat. Sci.* **6**, 209–224 (1991).
13. Guiot, J. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **80**, 49–69 (1990).
14. Guiot, J., Pons, A., De Beaulieu, J. L. & Reille, M. *Nature* **338**, 309–313 (1989).
15. Bartlein, P. J., Prentice, I. C. & Webb, T. *J. Biogeogr.* **13**, 35–57 (1986).
16. Huntley, B. *Quat. Proc.* **2**, 91–99 (1992).
17. Huntley, B. *Phil. Trans. R. Soc. B341*, 215–223 (1993).
18. Overpeck, J. T., Webb, T. & Prentice, I. C. *Quat. Res.* **23**, 87–108 (1985).
19. Guiot, J. *et al.* *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **103**, 73–93 (1993).
20. Huntley, B. *J. Quat. Sci.* (in the press).
21. Prentice, I. C., Bartlein, P. J. & Webb, T. *Ecology* **72**, 2038–2056 (1991).
22. Atkinson, T. C., Briffa, K. R. & Coope, G. R. *Nature* **325**, 587–593 (1987).
23. Huntley, B. & Prentice, I. C. in *Global Climates since the Last Glacial Maximum* (eds Wright, H. E. Jr. *et al.*) 136–168 (Univ. Minnesota Press, Minneapolis, 1993).
24. Hays, J. D., Imbrie, J. & Shackleton, N. *Science* **194**, 1121–1132 (1976).
25. Ruddiman, W. F. & McIntyre, A. *Quat. Res.* **16**, 125–134 (1981).
26. Broecker, W. S. & Denton, G. H. *Geochim. cosmochim. Acta* **53**, 2465–2501 (1989).
27. Weaver, A. J. & Hughes, T. M. C. *Nature* **367**, 447–450 (1994).
28. Prentice, I. C. *et al.* *J. Biogeogr.* **19**, 117–134 (1992).

ACKNOWLEDGEMENTS. We thank those who provided data and/or software used in our study, especially P. M. Anderson, P. J. Bartlein, S. Bottema, L. B. Brubaker, R. Cheddadi, W. Cramer, K. Gajewski, J. Guiot, M. E. Hutchinson, J. C. Ritchie and T. Webb. We thank also P. L. Gibbard, G. L. Jacobson Jr., M. Schopmeyer, C. Turner and R. G. West for helpful suggestions. The work was supported by the UK Natural Environment Research Council.

The role of soil organic matter in sustaining soil fertility

H. Tiessen*, E. Cuevas† & P. Chacon‡

* Department of Soil Science, University of Saskatchewan, Saskatoon S7N 0W0, Canada

† Centro de Ecología, Instituto Venezolano de Investigaciones Científicas, Caracas, Venezuela

‡ Universidad Nacional Experimental Simón Rodríguez, Caracas, Venezuela

MANY tropical soils are poor in inorganic nutrients and rely on the recycling of nutrients from soil organic matter to maintain fertility. In undisturbed rainforests such nutrients are recycled via the litter¹; 'slash-and-burn' agriculture, meanwhile, depends on the mineralization of organic nutrients from the plant remains^{2,3} or on (short-lived) inputs from ash⁴. This dependence on organic nutrients in tropical soils has the result that tests of soil quality which only give isolated measures of inorganic nutrient status are unreliable⁵, and that the effects of fertilization can be inconsistent because of leaching or fixation of inorganic nutrients. Here we attempt to quantify the role of organic matter in sustaining the fertility of soils from three different climate zones. We estimate rates of carbon turnover from ecological measurements and ¹⁴C dating, and determine its relation to the soil carbon and nutrient budgets. We find that agriculture without supplementary fertilization was economical for 65 years on temperate prairie and for six years in a tropical semi-arid thorn forest. An extremely nutrient-poor Amazonian soil showed no potential for agriculture beyond the three-year lifespan of the forest litter mat, once biological nutrient cycles were interrupted by slash-burning. These observations suggest that quantification of organic-matter cycling may provide an important guide to the agricultural potential of soils.

The contribution to soil fertility from the mineralization of organic matter (OM) can be estimated from budgets of carbon and nutrients during agricultural use. A more detailed picture of the mineralization process can be obtained when coarse and fine, mineral-associated fractions are differentiated^{6,7}, because turnover of clay- or silt-associated OM is slower than that of free organic debris⁸. As an alternative to the study of disturbed systems, turnover of litter and native OM are used in rainforests

to estimate organic contributions to the nutrient cycle^{9,10}. Plant litter enters soil OM with atmospheric ¹⁴C abundance, which subsequently declines with natural ¹⁴C decay. The mean residence time (MRT), and therefore the turnover of soil C, can thus be inferred from ¹⁴C abundance (radiocarbon dating) of a soil sample¹¹.

During cultivation, most mineralized organic N is lost and P reverts to unavailable forms, associated with Ca in temperate¹² or Fe and Al in tropical³ soils. In addition to the OM effects on N and P nutrition, reported here, OM contributes significantly to cation exchange (for example, Ca and K supply) and variable charge in tropical soils with low-activity clays. This may need to be considered to evaluate fully the fertility contributions of soil OM.

A sandy chernozemic prairie soil from the Canadian Great Plains, originally containing 8,750 g C m⁻² in the top 20 cm, lost 50% of its soil OM content during 65 yr of cultivation (Table 1 and ref. 6). The coarse OM fraction had accumulated substantial amounts of C in this cold, semi-arid grassland, where decomposition processes are slow. This coarse fraction accounted for approximately half of the total C loss, and was itself reduced by 70% (Table 1). The remaining 30% represent a new equilibrium pool, reflecting litter addition by the crop. Mineral-associated C was more stable, showing a mean residence time of 144 yr, calculated on the basis of the losses during cultivation. Mineral-associated C in the native soil showed a radiocarbon MRT of 330 yr, similar to other published values¹¹. Thus under cultivation, the MRT was halved from its native value because of the effects of tillage, changed moisture conditions and so on¹³.

A sandy ferralsol in semi-arid northeastern Brazil originally contained 60% less C than the Canadian soil (3,380 g C m⁻² in the top 20 cm), and lost 40% of its organic C in only 6 yr of cultivation (Table 1). The coarse C fraction, mainly decomposing roots of the native vegetation, was entirely lost but accounted for only 380 g m⁻² of the decrease. Most of the soil OM losses occurred from the mineral-associated fraction, which lost one-third of its C.

The ¹⁴C content of the silt-plus-clay fractions of the native soil was 113% of the modern value (113% modern), indicating the incorporation of post-1950 bomb ¹⁴C, and thus turnover times of <100 yr. A more precise date can be estimated using models that account for incorporation and decay of bomb ¹⁴C

TABLE 1 Changes in top-soil C on cultivation

	C content native (g m ⁻²)	C content disturbed (g m ⁻²)	C loss (g m ⁻²)	C loss (%)	MRT C loss* (yr)	MRT ¹⁴ C (yr)†
Canadian chernozem (prairie; 65 yr cultivation)						
Whole soil (–20 cm)	8,750‡	4,500‡	4,250‡	49‡	98‡	ND
Coarse C	3,250‡	1,000‡	2,250‡	70‡	55‡	ND
Mineral-assoc. C	5,500‡	3,500‡	2,000‡	36‡	144‡	330
Brazilian ferralsol (semi-arid thorn-forest; 6 yr cultivation)						
Whole soil (–20 cm)	3,380§	2,000§	1,380§	40§	11§	ND
Coarse C, root C	380§	0§	380§	100§	0.5§	ND
Mineral-assoc. C	3,000§	2,000§	1,000§	33§	15§	113% modern
Venezuelan ferralsol (rainforest; 3 yr slash-and-burn)						
Organic layer	7,000	1,300	5,700	81	2	ND
Whole soil (–15 cm)	5,100	3,600	1,500	29	9	ND
Coarse C, root C	2,800	ND	—	—	—	ND
Root C alone	500¶	500¶	0	0	0.4#	ND

* MRT was calculated as the inverse of the decomposition constant k , assuming that C declines from the native value C_n following $C = C_n e^{-kt}$. For the organic layer, k was obtained from litter decomposition studies and from the relationship of litter production to standing stock. For soils, k was based on three measurements of the C decline over three years for the Venezuelan, and one measurement each for the Canadian and Brazilian soils. †MRT equals radiocarbon age; ‡ ref. 6; § ref. 3; || ref. 19; ¶ refs 17, 18; # refs 10, 17; ND, not determined.

(ref. 14). Mineral-associated OM is likely to be a mixture of old and young portions^{13,14}, and the atmospheric ¹⁴C spike of the 1960s will have moved into and through portions of the OM to varying degrees. Using a simple two-compartment model¹³ the 113% modern indicates the presence of ~75% OM younger than 25 yr. Using a model based on ¹⁴C profiles with time after 1950 (which permits estimation of fast C turnover of samples taken since 1960¹⁵), one can ascribe an approximate MRT of 10–50 yr with a greater likelihood for the higher figure (compare with Fig. 1 in ref. 15). The MRT calculated from losses on cultivation is 15 yr (Table 1). The OM remaining after cultivation did not show an increased radiocarbon age, supporting findings that tropical soils may contain only a small recalcitrant C fraction¹⁶, with little effect on measured ¹⁴C age.

A sandy ferralsol of the Amazon rainforest at San Carlos de Rio Negro (Venezuela) was used for studies of OM turnover and nutrient cycling in the UNESCO-'Man and the Biosphere' programme^{10,17,19}, during which a site was slash-burned and analysed for 3 years, the normal duration of agriculture under locally practiced shifting cultivation. The slash-and-burn site had lost 81% of its litter layer and 29% of its soil C to 15 cm depth over 3 years (Table 1), giving MRTs of 2 and 9 yr, respectively, under the experimental (uncropped but weedy) conditions.

We sampled the same soil in 1991 at an adjacent forest site, which contained 7,000 g C m⁻² in the litter layer and root mat (Table 2). This mat was entirely above the mineral soil surface, and is responsible for a close recycling of nutrients in the intact forest^{1,20}. The first 15 cm of mineral soil contained a further 5,500 g C m⁻² at a concentration of 25 mg g⁻¹. Total C stores were thus 40% higher than in the prairie soil. Between 15 and 39 cm depth in the mineral soil, a further 6,000 g C m⁻² were found at a concentration of 23 mg g⁻¹ in the soil fines and 13 mg g⁻¹ in lateritic nodules, which made up 60% of the horizon weight.

Mineral-associated C was 112% modern ¹⁴C in the top soil, and 105% modern between 36 and 60 cm depth, which translates into MRTs of 30–50 yr and close to 100 yr respectively, using the models and data of Harkness *et al.*¹⁵ (Table 2). In the 36–60 cm layer, ¹⁴C of the coarse OM fraction was 172% modern. Peak ¹⁴C levels of 165% modern were recorded in wood cores from Amazonian forests for 1965²¹, and it seems that the coarse organic matter found in the mineral soil preserved this peak without dilution 26 yr after the atmospheric peak. This may be the result of leaf- and litter-harvesting activities of leafcutter ants and termites which construct an average of 1,500 nests per hectare in this ferralsol²². Similarly high ¹⁴C abundances have been observed in other forest soils (A. F. Harrison, personal communication). Carbon contained in the lateritic nodules,

which comprise most of that soil layer and contain 44% of its C, had a very long MRT of 4,600 yr (Table 2).

In the prairie soil, organic-matter mineralization was accompanied by the liberation of 40% of the organic N and P reserves⁶, a loss of 360 g N m⁻² and 39 g P m⁻². Labile inorganic P reserves were reduced by 22% (20 g P m⁻²)¹². Both cultivation and organic matter mineralization continued beyond 65 yr, but the reduced amounts and rates of organic matter mineralization caused a decrease in P and N supply and necessitated fertilization for economical crop production. The trend of increased fertilizer additions during the period 60–80 yr after the land was broken is typical for the prairie region and contributes to the difficult economic situation of Great Plains agriculture in the 1990s.

The Brazilian thorn-forest site was abandoned after 6 yr of cultivation because of low productivity. As inorganic nutrient availability is very low in both the native and cultivated soil, the loss of fertility may be attributed to lowered nutrient release due to a decrease of 30% in soil organic N and P (100 and 1.9 g m⁻² respectively)³. Labile organic and inorganic P converted to low-availability inorganic P at 2.4 g m⁻² (ref. 4). As in the prairie soil, organically held nutrient stocks have been depleted below levels adequate for unfertilized agriculture after 40–50% of the OM were lost, that is, after about half the MRT of the soil OM under cultivation, or about one-fifth of the native MRT of the mineral-associated C. Fertilizer additions are not an economic option for most subsistence farmers of northeast Brazil, and the normal practice is abandonment of cultivated sites to bush fallow or pasture.

In the undisturbed Venezuelan rainforest, an OM budget to 60 cm depth shows the following C turnover and associated nutrient contribution: 60% of the C, 65% of the N and 50% of the P are in particulate OM with an MRT of less than 4 yr; 27% of the C, 29% of the N and 33% of the organic P are mineral-associated and have native MRTs with an average near 50 yr. This is similar to the mineral-associated OM in the semi-arid ferralsol from northeast Brazil. The remainder of the OM and associated nutrients are sequestered in lateritic nodules and do not participate in nutrient cycles. These natural rates of nutrient cycling will be accelerated under agriculture, at least by the factors observed at the two other sites, that is, over half the nutrients will mineralize in ~2 yr. The remainder will contribute small amounts of nutrients for a further 25 yr. Comparisons with the other ecosystems, and experience of local shifting cultivators, indicate that the rates of nutrient release at this point are not sufficient for agricultural crops. Thus agriculture is not sustainable without nutrient inputs beyond 3 yr, although the release of remaining nutrients can provide for the re-establishment of secondary successions. These often show higher net productivity

TABLE 2 Organic-matter stock of soil and litter layer under natural rainforest at San Carlos

Sample, depth	Subsample	Carbon (g m ⁻²)	Nitrogen (g m ⁻²)	Organic P (g m ⁻²)	MRT (yr)
Fresh litter, 0–1 cm	Total	65	1.6	0.02	1.5*
Org. layer, 1–21 cm	Total	6,930	315	12.1†	1.5–3.3*
Min. soil, 21–36 cm	Total	5,145	239	7.7	ND
	Coarse C	2,835	116	4.9	ND
	Clay + silt	2,310	123	2.8	30–50‡§
Min. soil, 36–60 cm	Total	5,799	295	9.6	ND
	Soil <0.5 mm	3,274	246	6.2	ND
	Coarse C	696	118	1.4	mod‡§
	Clay + silt	2,578	128	4.8	~100‡§
	Nodules >0.5 mm	2,525	49	3.4	4,600‡
Total to 60 cm		17,939	850	22.7	

* Calculated from refs 9 and 10. † Includes 6.7 g m⁻² of inorganic P associated with the organic matter. ‡ Based on radiocarbon dates. § These materials date 'modern', that is, <100 yr. A more precise interpretation has been attempted based on the dynamics of hydrogen-bomb ¹⁴C; see text.

than the mature native forest and can re-establish relatively high soil OM levels¹⁹, although basal area and total biomass of the mature forest are reached only after 190 yr (ref. 23) when internal nutrient cycles are re-established. Nutrient uptake by forest, agriculture or regrowth from below the sample depth is unlikely in the San Carlos soil because practically no roots are found below 40 cm (ref. 17), where root growth is limited by Ca and P deficiency²⁴.

Given the high rainfall of the region (>3,500 mm yr⁻¹) and limited stabilization of organic and other materials by the silts and clays (10% of soil weight), fertilization is unlikely to be an option even for cash crops, unless a successful litter-organic matter cycle can be established. Agroforestry aims at such OM management, using short-lived litter mulches from trees to release nutrients to crops.

In the less leached northeast Brazilian soil, fertilization might be an option, despite similarly low silt and clay contents, but potential crop failures because of drought limit fertilizer investments. The short MRT of the remaining OM predicts continuing rapid degradation, making OM management essential. After 6 yr of use, forest regrowth takes >50 yr, although soil OM reaches near-native levels within 10 yr of bush fallow⁶.

The identification of tropical rainforest soils suitable for agriculture has been under some discussion. On one hand it has been shown that with suitable management and inputs, agricultural production can be sustained at economically viable levels on finer textured soils of the Amazon²⁵. On the other there is the opinion that much of the region might best be managed without deforestation²⁶. Rates of nutrient cycling associated with soil OM turnover under primary or secondary forest may provide a predictive tool for evaluating the potential of soils for agriculture and for subsequent forest recovery. An important caution against the introduction of arable agriculture is that under cultivation, C loss was twice as rapid as the C mineralization associated with OM turnover in the natural systems, and therefore the potential for nutrient loss is greatly enhanced. The development of easily measured indicators for nutrient cycling, and the evaluation of nutrient cycling and OM turnover in natural and cultivated soils, will be essential for decisions on land use and management where measurements of static nutrient pools are inadequate. An understanding of OM cycling can also be applied to agroforestry, which aims to maintain biological nutrient cycles with low external inputs. □

Received 1 November 1993; accepted 23 September 1994.

- Medina, E. & Cuevas, E. in *Mineral Nutrients in Tropical Forest and Savanna Ecosystems* (ed. Proctor, J.) 217–240 (Blackwell Scientific, Oxford, 1989).
- Mueller-Harvey, I., Juo, A. S. R. & Wild, A. J. *Soil Sci.* **36**, 585–591 (1985).
- Tiessen, H., Salcedo, I. H. & Sampaio, E. V. S. B. *Agric. Ecosyst. Envir.* **38**, 139–151 (1992).
- Stromgaard, P. *Pl. Soil* **80**, 307–320 (1984).
- Adepetu, J. A. & Corey, R. B. *Soil Sci.* **122**, 159–164 (1976).
- Tiessen, H. & Stewart, J. W. B. *Soil Sci. Soc. Am. J.* **47**, 509–514 (1983).
- Vitorello, V. A., Cerri, C. C., Andreux, F., Feller, C. & Victória, R. L. *Soil Sci. Soc. Am. J.* **53**, 773–778 (1989).

- Feller, C., Guiraud, G., Hetier, J. M. & Morel, C. *Int. J. trop. Agric.* **1**, 123–130 (1983).
- Cuevas, E. & Medina, E. *Oecologia* **68**, 466–472 (1986).
- Cuevas, E. & Medina, E. *Oecologia* **76**, 222–235 (1988).
- Martel, Y. A. & Paul, E. A. *Can. J. Soil Sci.* **54**, 419–426 (1974).
- Tiessen, H., Stewart, J. W. B. & Moir, J. O. *J. Soil Sci.* **34**, 815–823 (1984).
- Harrison, K. G., Broecker, W. S. & Bonani, G. *Science* **262**, 725–726 (1993).
- O'Brien, J. *Soil Biol. Biochem.* **16**, 115–120 (1984).
- Harkness, D. D., Harrison, A. F. & Bacon, P. J. *Radiocarbon* **28**, 328–337 (1986).
- Trumbore, S. E. *Glob. Biogeochem. Cycles* **7**, 275–290 (1993).
- Sanford, R. L. *J. trop. For. Sci.* **1**, 268–279 (1989).
- Stark, N. & Spratt, M. *Trop. Ecol.* **18**, 1–9 (1977).
- Uhl, C. *J. Ecol.* **75**, 377–407 (1987).
- Stark, N. & Jordan, C. *Ecology* **59**, 434–437 (1978).
- Worbes, M. & Junk, W. J. *Ecology* **70**, 503–507 (1989).
- Saleck, J., Herrera, R. & Jordan, C. F. *Biotropica* **15**, 1–7 (1983).
- Saldarriaga, J. G., West, D. C., Tharp, M. L. & Uhl, C. *J. Ecol.* **76**, 938–958 (1988).
- Cuevas, E. & Medina, E. in *Root Ecology and its Practical Application* (eds Böhm, W., Kutschera, L. & Lichtenegger, E.) 653–666 (Bundesanstalt für Alpenländische Landwirtschaft, Irdning, Austria, 1983).
- Sanchez, P. A. & Benites, J. R. *Science* **238**, 1521–1527 (1987).
- Fernside, P. *Bioscience* **37**, 209–214 (1987).

Subduction zone melting of pelagic sediments constrained by melting experiments

Geoff T. Nichols*, Peter J. Wyllie* & Charles R. Stern†

* California Institute of Technology, Division of Geological and Planetary Sciences, Mail Code 170-25, Pasadena, California 91125, USA

† University of Colorado, Department of Geological Sciences, Campus Box 250, Boulder, Colorado 80309, USA

THE fate of pelagic sediments in subduction zones has an important bearing on global geochemical cycles and the thermal structure of subduction zones. Recent mass-balance calculations¹ have indicated that about 20% of subducted sediments are recycled to volcanic arcs. Trace-element studies² of these volcanic arcs imply that the subducted sediment melts while the gabbroic crust that underlies it dehydrates. This requires a rather specific thermal structure in subduction zones. Here we report laboratory melting experiments that allow us to derive melting curves for pelagic red clay at pressures of up to 40 kbar, equivalent to depths of 120 km within the mantle. The melting behaviour of wet red clay is similar to that for wet gabbro, but melting occurs at a slightly lower temperature, showing that sediment melting and gabbro dehydration can occur at the same temperature. The combination of trace-element data and phase diagrams such as that derived here may thus be used to constrain the temperature of the slab–mantle boundary.

The mass contribution from subducted oceanic crust to volcanic arcs has been debated since Green and Ringwood³ first demonstrated experimentally that the partial melting of sub-

ducted oceanic crust beneath volcanic arcs could generate magmas approximating the composition of andesites. By the 1980s^{4,5} the prevailing opinion was that arc magmas were generated by melting in the mantle wedge, with geochemical signatures from oceanic crust transferred by fluids of uncertain source and character. According to recent petrological and geochemical studies some magmas (for example, adakites) are derived by partial melting of subducted gabbroic crust^{6,7}, but it is agreed that this requires special circumstances such as young, hot lithosphere with slow subduction⁸.

The contribution of melts from subducted sediments to arc magmas was proposed in the 1960s^{9,11}, and subsequently disputed^{12,13}. The involvement of sediments in arc magmas is now confirmed by evidence from Sr, Pb and Be isotopes^{14,17} and trace elements^{15,18,19}. Most pelagic sediments are subducted²⁰. They mostly consist of combinations of terrigenous components (red clay), biogenic phases (opal and calcite), and about 33% water²¹. We have new experimental data on a pelagic red clay that, combined with earlier data on siliceous limestones, define the behaviour of pelagic sediments.

The starting material is an abyssal red clay from the Atlantic Ocean. The assemblage of quartz, muscovite, plagioclase, illite, halite with minor carbonate and kaolinite contained volatile components: 7.79% H₂O⁺, 1.05% H₂O⁻, 1.7% CO₂, 1.9% Cl and 0.06% S. Our earlier study of this sediment was based on X-ray powder diffraction (XRD) and optical studies of crushed fragments^{22,23}, but examination of our new experiments with scanning electron microscope (SEM) revealed complexities. Experiments were conducted in piston-cylinder apparatus, with earlier short-duration (1–16 hour) runs in Ag₃₀Pd₇₀ capsules, and all recent long-duration (155–207 hour) runs in Au capsules. Experiments were conducted with the sample dry, and with about 20% H₂O added (8% and 28% H₂O experiments, respectively).

Partly dehydrated subsolidus experiments formed the assemblages shown in Fig. 1, with much original H₂O displaced to the NaCl-rich vapour phase. Near-solidus runs were dominated by <5 µm mica crystals. The solidus was located by identification of low refractive index glass in crushed fragments, and the presence of narrow (1–4 µm) rims of glass around minerals (see Fig. 2a) in two of the three runs at 650 °C. The presence of significant

proportions of liquid at 650 °C between 12 and 20 kbar was demonstrated by the sinking of <15 µm quartz into the lower half of the capsule (Fig. 2a). The liquid precipitated abundant mica crystals during the quench.

Figure 2c–e shows definitive evidence for the precipitation of quench minerals from liquid. Figure 2c is a diagrammatic representation of runs similar to those in Fig. 2d–e, showing the sample with a border of glass (100–150 µm wide) surrounding a central core of glass plus minerals, with a few large biotite crystals settled from liquid during the run. Quench crystals nucleated and grew only in the more slowly cooled central part of the capsule. The enlarged view in Fig. 2e, of a superliquidus run, shows that garnets (5–10 µm) have atoll-like skeletal forms, characteristic of rapid growth. Toward the more slowly cooled centre of the charge these atoll forms are replaced by solid, equant quenched garnets. In the centre 5 µm, acicular rosettes of quenched kyanite also occur with rare grains of rutile, along with quenched biotite in the lower temperature runs. Figure 2a and b shows small (~5 µm) garnets in the top halves of charges. Garnet should sink to the bottom of the capsule more efficiently than did the less dense quartz; the garnet grew during the quench. The matrix of fine-grained quenched mica dominated the charges through an interval of 100 °C above the solidus.

The liquidus boundary above 10 kbar in Fig. 1 corresponds to the limit of biotite. Primary biotite was identified in 750 °C experiments where it grew to 50 µm, and at 800 °C biotite up to 180 µm was observed, but in runs below 750 °C we have not been able to distinguish between primary and quench micas. Garnet, clinopyroxene, and amphibole should occur as primary minerals above the solidus, but we have found no criteria to

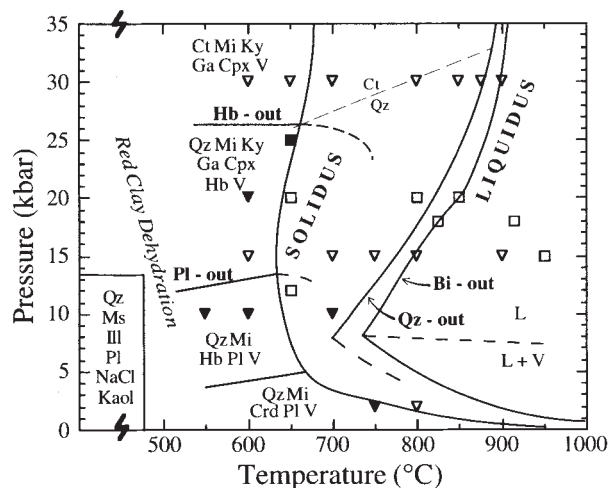


FIG. 1 *P*–*T* diagram showing melting relations of red clay, collected in the Atlantic (24° 30' N, 64° 47' W) and provided by C. Emiliani. Black and white symbols indicate experiments with 28% and 8% water, respectively. Long duration runs indicated with squares, and shorter runs with inverted triangles. Mineral abbreviations: Bi, biotite; Cpx, clinopyroxene; Crd, cordierite; Ct, coesite; Ga, garnet; Hb, hornblende; Ill, illite; Kaol, kaolinite; Ky, kyanite; L, liquid; Mi, mica; Ms, muscovite; Pl, plagioclase; Qz, quartz; V, vapour. The boundary between L and L+V occurs at the pressure where 8% H₂O saturates granitoid liquids³⁷.

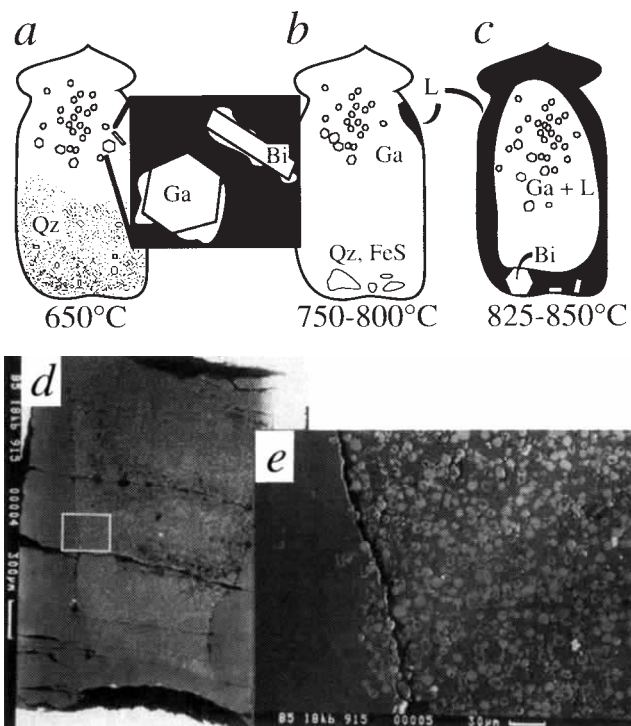
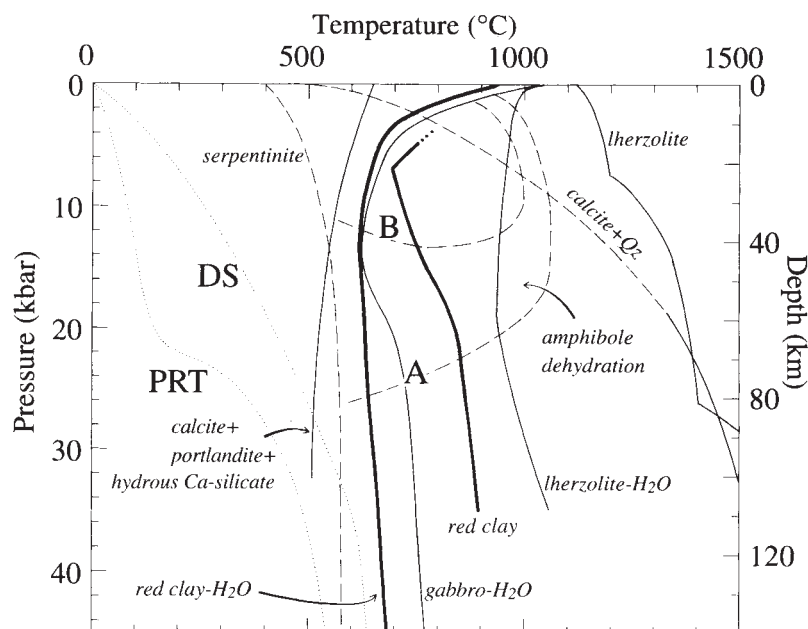


FIG. 2 a–c, Schematic diagrams depicting the mineral changes through the melting interval at pressures of 18–20 kbar: a, experiments at 650 °C, immediately above the solidus, have narrow glass rims on quenched minerals, and quartz sinks to the lower half of the capsules; b, between 750–800 °C the volume of melt increases, and quenches as glass margins on the capsules; larger quartz grains occur on the lower capsule margins; c, glass margins in experiments between 825–850 °C are much larger, and quenched garnet occurs in the centre; equilibrium assemblage is biotite+liquid. d, Superliquidus run at 915 °C with thick marginal glass. e, An enlargement through the transition from glass, to quenched skeletal garnets, and to euhedral garnets in the centre.

FIG. 3 P - T diagram comparing solidi and dissociation curves (dashed) of subduction zone lithologies with model P - T trajectories for the top of subducting oceanic crust from Davies and Stevenson³⁰ (DS), (30° dipping slab, with velocity 7.2 cm yr⁻¹, no shear heating) and Peacock *et al.*⁸ (PRT) (26.6° dipping slab, with velocity 10 cm yr⁻¹, no shear heating). Two bold solidus curves for subducted red clay are from Fig. 1; lower temperature solidus is clay with pore fluid, and the higher temperature solidus is the deduced vapour-absent melting based on the biotite-out curve in Fig. 1 (vapour present liquidus). Dehydration melting of amphibolite occurs between B and A, releasing H₂O. Curves from the following references: serpentinite³⁷, calcite + portlandite + hydrous Ca-silicate³⁸, gabbro-H₂O³⁵, amphibole dehydration³⁵, calcite + quartz²⁹, lherzolite³⁹ and lherzolite-H₂O⁴⁰.



distinguish them from quench minerals. Quartz is primary (Fig. 2a, b) with no indications that it grows during quench. The quartz-out boundary in Fig. 1 is based on the disappearance of quartz from XRD patterns of crushed runs and from SEM scans.

The solidus is almost identical within experimental limits with that for granite-H₂O^{24,25}, thus it was drawn corresponding to the granite-H₂O solidus in the interval 2–10 kbar. Johnson and Plank²⁶ reported a solidus temperature (T_s) for a similar red clay between 800 °C and 850 °C at 30 kbar, corresponding to the biotite-out curve and liquidus in our sediment (Fig. 1). The disparity in the experimentally determined T_s may be explained by normative calculations which estimate ~1% Qz for Johnson and Plank's²⁶ clay, compared with ~15% Qz for our red clay. This indicates that although solidi in both sets of experiments should be identical, the percentage of liquid in Johnson and Plank's experiments may be too low to detect until mica begins to dehydrate.

The sinking of quartz at 650 °C and 700 °C (Fig. 2a) suggests that tens of per cent of liquid are produced within ~50 °C of the solidus for the 8% sediment, with ~5–7% H₂O in subsolidus NaCl-rich vapour. Tens of per cent of liquid may also be expected within a few degrees of the solidus for dehydration-melting of the sediment at higher temperatures, near the biotite-out curve (Fig. 1)²⁷. Microprobe analyses indicate peraluminous siliceous glasses²⁸. Considering the composition changes that would be caused by the known quench minerals, micas, garnet and kyanite, we conclude that the liquid compositions are broadly granitoid, as expected.

Figure 3 compares melting curves relevant for subducted materials with P - T paths for the surface of subducted oceanic lithosphere. Pelagic sediments with aqueous pore fluid begin to melt at temperatures lower than gabbroic crust (compare red clay-H₂O and gabbro-H₂O curves). In contrast, decarbonation of pelagic limestone occurs only at very high temperatures (compare calcite + Qz and data from Davies and Stevenson³⁰, denoted by DS in Fig. 3), unless H₂O is present²⁹. With sufficient H₂O, siliceous limestone becomes hydrated and partial melting begins at the calcite + portlandite + hydrous Ca-silicate curve, at even lower temperatures than for red clays, producing carbonatitic liquids. Decarbonation and partial melting reactions in subducted limestones are probably localized by the passage of aqueous pore fluids along channels, with impermeable masses of limestone untouched by the fluid and transported to deeper levels without reaction (see the calcite + Qz curve).

Many computed thermal models for subduction zones have yielded high-temperature P - T paths that required significant melting of subducted sediments and gabbroic crust at moderate depths. Davies and Stevenson³⁰ critically reviewed previous models, concluding that the thermal field was more constrained than had been assumed. They showed that it is difficult to heat the slab-mantle interface, and although this region was poorly modelled in their calculations, they were confident with their conclusion that the steady-state subducted plate is not melted extensively, if at all. The P - T path DS (Fig. 3), one of their calculated paths for the top of a subducted slab, remains below the solidus curves for red clay and gabbro/eclogite to at least 140 km, and at deeper levels apparently subparallels the red clay solidus. The numerical models of Peacock, Rushmer and Thompson⁸ also predicted cold P - T trajectories for subducted oceanic crust (denoted by PRT in Fig. 3) with no partial melting, except under rare circumstances such as high rates of shear heating or slow subduction of young, hot oceanic lithosphere. Davies and Stevenson³⁰, contrary to Molnar and England³¹, maintained that frictional heating is not a major heat source. In this analysis it is recognized that higher temperature P - T trajectories are possible under special conditions, however, cooler P - T paths without shear heating were selected following current model estimates of steady-state subduction conditions.

After compaction and dewatering of sediments beneath the forearc, the oceanic crust passes through the blueschist and into the eclogite facies experiencing a series of dehydration reactions⁸. Our results (Fig. 1) indicate the partly dehydrated subsolidus sediment, at depth, consists of eclogitic garnet and clinopyroxene, with mica, quartz/coesite and kyanite, although some sediment components would have been leached by the escape or passage of aqueous vapour. Melting reactions calcite + portlandite + hydrous Ca-silicate, red clay-H₂O and gabbro-H₂O (Fig. 3) require aqueous vapour, available from serpentinite within the gabbroic crust³². Dehydration of serpentinite begins in the upper part of the slab at depths shallower than the intersection depth of P - T paths with the red clay-H₂O solidus (although in the lower slab dehydration is delayed to greater depths). H₂O is also released from amphibole between depths of about 35–75 km (B→A, Fig. 3), with maximum dehydration generally considered to occur near A^{8,33–35}.

According to the thermal structures DS and PRT, neither red clay nor gabbroic crust beings to melt at depths associated with overlying arc volcanism. Extrapolation of the gabbro-H₂O soli-

dus and red clay-H₂O solidus makes intersection appear unlikely at any depth. However, only a slight temperature increase in the P - T path DS would be sufficient to cause intersection with the red clay-H₂O solidus, placing the P - T path between the red clay-H₂O curve and the gabbro-H₂O solidus. This produces a situation where through a considerable depth interval the sediments begin to melt, but no melting occurs within the gabbroic crust.

This is precisely the situation proposed by Plank and Langmuir³⁶ whose conclusions, based on geochemical mass balances for various trace elements, are satisfied by the experimentally determined solidus curves and a P - T trajectory for the top of the subducted slab somewhere between the red clay-H₂O and gabbro-H₂O curves in Fig. 3. After extraction of liquid the sediment contains residual biotite and quartz/coesite, and we expect garnet and kyanite to coexist as well at pressures above 15–20 kbar (with amphibole through a limited range). Our trace-element analyses indicate residual biotite depletes the melt in Ba, K and Ti by factors ~5, ~3 and ~9 times, respectively.

This experimental demonstration of the feasibility of Plank and Langmuir's³⁶ deduction that, in many subduction zones, sediments melt while gabbro does not is encouraging. If confirmed, it places specific constraints on the temperatures at the slab-mantle interface, indicating that temperatures remain between the solidi of red clay-H₂O and gabbro-H₂O, ~650–750 °C, through a significant depth interval. Additional correlation between petrology, geochemistry, experimental petrology and thermal modelling is required if the flux of crustal material into subduction zones, and its redistribution between arc magmas, lithosphere and deeper reservoirs, is to be worked out in detail. □

Received 30 June; accepted 14 September 1994.

- Plank, T. & Langmuir, C. H. *Nature* **362**, 739–742 (1993).
- Plank, T. thesis, Columbia Univ. (1993).
- Green, T. H. & Ringwood, A. E. *Contr. Miner. Petrol.* **18**, 105–162 (1968).
- Basaltic Volcanism Study Project *Basaltic Volcanism on the Terrestrial Planets* (Pergamon, New York, 1981).
- Gill, J. *Orogenic Andesites and Plate Tectonics* (Springer, Heidelberg, 1981).
- Defant, M. J. & Drummond, M. S. *Nature* **347**, 662–665 (1990).
- Kay, S. M., Ramos, V. A. & Marquez, M. J. *Geol.* **101**, 703–714 (1993).
- Peacock, S. M., Rushmer, T. & Thompson, A. B. *Earth planet. Sci. Lett.* **121**, 227–244 (1994).
- Coats, R. R. in *The Crust of the Pacific Basin* 92–109 (Geophysics Monograph 6, American Geophysical Union, Washington DC, 1962).
- Armstrong, R. L. *Rev. Geophys.* **6**, 175–199 (1968).
- Oxburgh, E. R. & Turcotte, D. L. *Nature* **218**, 1041–1043 (1968).
- Sinha, A. K. & Hart, S. R. *Carnegie Inst. Wash. Yearb.* **71**, 309–312 (1972).
- Oversby, V. M. & Ewart, A. *Contr. Miner. Petrol.* **37**, 181–210 (1972).
- Kay, R. W., Sun, S.-S. & Lee-Hu, C.-N. *Geochim. cosmochim. Acta* **42**, 263–273 (1978).
- White, W. M. & Dupré, B. J. *geophys. Res.* **91**, 5927–5941 (1986).
- Tera, F. et al. *Geochim. cosmochim. Acta* **50**, 535–550 (1986).
- Morris, J. D., Leeman, W. P. & Tera, F. *Nature* **344**, 31–36 (1990).
- Hawkesworth, C. J., Hergt, J. M., Ellam, R. M. & McDermott, F. *Phil. Trans. R. Soc.* **335**, 393–405 (1991).
- Hawkesworth, C. J., Gallagher, K., Hergt, J. M. & McDermott, F. *Phil. Trans. R. Soc.* **342**, 179–191 (1993).
- Moore, J. C. *Geology* **3**, 530–532 (1975).
- Rea, D. K. & Ruff, L. J. *SUBCON* (abstr.) 51 (1994).
- Stern, C. R. thesis, Univ. Chicago (1973).
- Stern, C. R. & Wyllie, P. J. *Contr. Miner. Petrol.* **42**, 313–323 (1973).
- Boettcher, A. L. & Wyllie, P. J. *J. Geol.* **76**, 235–244 (1968).
- Stern, C. R. & Wyllie, P. J. *J. geophys. Res.* **86**, 10412–10422 (1981).
- Johnson, M. C. & Plank, T. *Eos* **74**, 680 (1993).
- Vielzeuf, D. & Holloway, J. R. *Contr. Miner. Petrol.* **98**, 257–276 (1988).
- Nichols, G. T., Wyllie, P. J. & Stern, C. R. *SUBCON* (abstr.) 208–210 (1994).
- Huang, W.-L., Wyllie, P. J. & Nehru, C. E. *Am. Miner.* **65**, 285–301 (1980).
- Davies, J. H. & Stevenson, D. J. *J. geophys. Res.* **97**, 2037–2070 (1992).
- Molnar, P. & England, P. J. *J. geophys. Res.* **95**, 4833–4856 (1990).
- Ringwood, A. E. *Chem. Geol.* **82**, 187–207 (1990).
- Lambert, I. B. & Wyllie, P. J. *J. Geology* **80**, 693–708 (1972).
- Pawley, A. R. & Holloway, J. R. *Science* **260**, 664–667 (1993).
- Wyllie, P. J. & Wolf, M. B. in *Magmatic Processes and Plate Tectonics* (eds Prichard, H. M., Alabaster, T., Harris, N. B. W. & Neary, C. R.) 405–416 (Spec. Publ. 76, Geol. Soc., London, 1993).
- Plank, T. & Langmuir, C. H. *Eos* **73**, 637 (1992).
- Wyllie, P. J. *Am. Miner.* **64**, 469–500 (1979).
- Boettcher, A. L. & Wyllie, P. J. *Geochim. cosmochim. Acta* **33**, 611–632 (1969).
- Takahashi, E. J. *geophys. Res.* **91**, 9367–9382 (1986).
- Green, D. H. *Earth planet. Sci. Lett.* **19**, 37–53 (1973).

ACKNOWLEDGEMENTS. We thank A. Pawley and S. Peacock for constructive reviews. Research supported by the National Science Foundation.

How long should flowers live?

Tia-Lynn Ashman* & Daniel J. Schoen

Biology Department, McGill University, 1205 Dr. Penfield Avenue, Montréal, Québec H3A 1B1, Canada

FLORAL longevity, the length of time a flower remains open and functional, varies among plant species. Flowers of some species live less than one day (morning glory), whereas others live for several weeks (orchids)^{1–3}. By viewing floral longevity as a resource allocation strategy^{2,4}, we now incorporate the study of its evolution into the well developed theoretical framework provided by evolutionarily stable strategy models that address variation in life history^{5,6}. Flowers must remain open to contribute to plant fitness through ovule fertilization and pollen dissemination, when they require resources for respiratory maintenance and pollinator attraction. Accordingly, floral senescence should occur when the expected fitness gain per unit of floral maintenance investment diminishes to the point where it becomes more profitable to construct a new flower than to maintain an existing one. Our experimental evidence supports floral longevity as an adaptation that balances rates of pollen receipt and removal against the cost of floral maintenance.

The floral longevity model⁴ (Fig. 1) tested here assumes that there is a fixed pool of resources for flowering and hence a trade-off between the number of flowers made and the construction and maintenance of each flower. It also assumes that variation in floral longevity is heritable and can be optimized by natural selection. Optimal floral longevity is determined by the interaction of three major factors: (1) the daily cost of maintaining a flower relative to the cost of constructing a new flower; (2) the rate at which pollen is received to fertilize ovules ('female fitness accrual rate'); and (3) the rate at which pollen is disseminated and enters the pool of pollen that competes to fertilize ovules ('male fitness accrual rate') (Fig. 1). Constant daily rates of female and male fitness accrual are assumed such that a flower's remaining contribution to plant fitness diminishes with time; that is, the proportion of pollen that is undisseminated and the proportion of ovules that are unfertilized approaches zero. Different combinations of these parameters result in different optimal floral longevity of $t^* = 1, 2, \dots T$ days (Fig. 1).

To determine how well this theoretical framework accounts for variation in floral longevity, we collected data on floral maintenance costs, male and female fitness accrual rates, and floral longevity for eleven species (representing ten families) growing in their native habitats (Table 1; Fig. 2). Ten of these species were studied in the Rocky Mountains, Colorado, and the eleventh (*Trillium grandiflorum*) was studied at Mont St Hilaire, Québec. The daily rates at which female and male fitness accrued varied among species (Fig. 2; Table 1): they ranged from rapid, as in *Oenothera flava*, in which 82–97% of male and female fitness was accrued in a single day, to slow, as in *Trillium grandiflorum*, in which only 6% and 17% of male and female fitness was accrued each day. The daily cost of maintaining flowers ranged from 2% (*Sedum lanceolatum*) to 9% (*Geranium caespitosum*) of the cost of constructing a new flower (Table 1). These empirically derived values of male and female fitness accrual, and flower maintenance cost, were substituted into the floral longevity model⁴ (Fig. 1), which was then used to obtain the optimal longevity for each species (t^*). The optimal floral longevity was compared to floral longevity observed in the field. Floral longevity measured in the field has been shown to be a good measure of the maximum floral longevity in the absence of pollinators³.

* Present address: Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA.

The correlation between observed and predicted optimal floral longevity was positive and highly significant (Spearman's rank correlation, $r_s = 0.68$; $P = 0.02$; $N = 11$) (Fig. 3). This correlation is not driven by variation in protandry (correlation between protandry and floral longevity: $r_s = 0.40$; $P > 0.2$; $N = 11$), nor

by the number of flowers produced ($r_s = 0.19$; $P > 0.5$; $N = 11$), but these characters do have explicit roles in the model formulation⁴. Thus, there is good qualitative agreement between predicted and observed floral longevity of the eleven species studied. A close quantitative fit, however, was not found. When

TABLE 1 Characteristics of species investigated for optimal floral longevity

Species	Habitat: pollinator: breeding system	Ratio (m) of floral maintenance to construction cost (s.e.)	Daily fitness accrual rates (s.e.)	
			Male P(1)	Female G(1)
<i>Aquilegia caerulea</i>	Shady, aspen groves (M); hummingbirds and bumblebees; self-compatible ¹³ , protandrous	0.07 (0.00)	0.21 (0.01)	0.19 (0.03)
<i>Cerastium arvense</i>	Exposed, open slopes (A); Small flies; unknown	0.01 (0.00)	0.55 (0.04)	0.17 (0.02)
<i>Geranium caespitosum</i>	Sandy, dry slopes (M); large bees and bumblebees; unknown	0.09 (0.01)	0.86 (0.03)	0.50 (0.04)
<i>Lewisia pygmaea</i>	Moist gravel slopes (A); flies?; self-compatible*	0.08 (0.02)	0.73 (0.02)	0.94 (0.04)
<i>Mertensia ciliata</i>	Near willow shrubs (A); syrphid flies; self-compatible ¹⁴	0.07 (0.01)	0.82 (0.02)	0.53 (0.04)
<i>Oenothera flava</i>	Dry, rocky slopes (M); hawkmoths; self-compatible?	0.02 (0.01)	0.82 (0.05)	0.97 (0.03)
<i>Pedicularis sudetica</i>	Wet to moist meadows (M); bumblebees; unknown	0.04 (0.01)	0.26 (0.03)	0.05 (0.01)
<i>Polemonium delicatum</i>	Near willow shrubs (A); syrphid flies and small flies self-compatible*	0.04 (0.01)	0.60 (0.05)	0.55 (0.05)
<i>Sedum lanceolatum</i>	Dry, exposed slopes (A); ants, flies; self-compatible*	0.02 (0.01)	0.55 (0.05)	0.07 (0.01)
<i>Trillium grandiflorum</i>	Beech-maple forest (T); bumblebee?; self-compatible?†	0.02 (unavailable)	0.06 (0.01)	0.17 (0.01)
<i>Zigadenus elegans</i>	Shady, aspen groves (M); small flies, gnats; self-compatible*	0.08 (0.01)	0.38 (0.02)	0.07 (0.01)

Habitat, pollinator type and protandry were noted in the field. Codes for habitat zones are as follows: A, alpine; M, montane; T, temperate forest. Floral maintenance cost for each species is expressed relative to construction costs. Daily maintenance costs (in mg carbon) of nectar production and corolla respiration were calculated for each species. Nectar production for each species was determined for 4–13 flowers protected from pollinators. Nectar present in flowers at anthesis was considered a construction cost, and sugar carbon content was added to corolla construction costs. Nectar produced after anthesis was considered a maintenance cost. Carbon liberated through corolla respiration was measured directly using a LICOR (no. 6250) portable photosynthesis system (as in the case of *Trillium*) or estimated for the smaller flowers which could not be directly measured: these estimates were based on the species-specific corolla biomass measures and the mean respiration rate previously measured on two large flowered species (*Trillium* and *Clarkia*: 0.0133 mg carbon lost per day per gram corolla tissue). The amount of carbon required to construct a corolla of a given biomass was determined from data on composition of and carbon cost per chemical constituent of corolla tissue^{15–18} and 4–13 measures of corolla biomass per species. Fitness accrual rates were calculated from the functions described in Fig. 2.

* C. Galen, personal communication.

† Related species are self-compatible¹⁹ and pollen deposition limits seed-set²⁰.

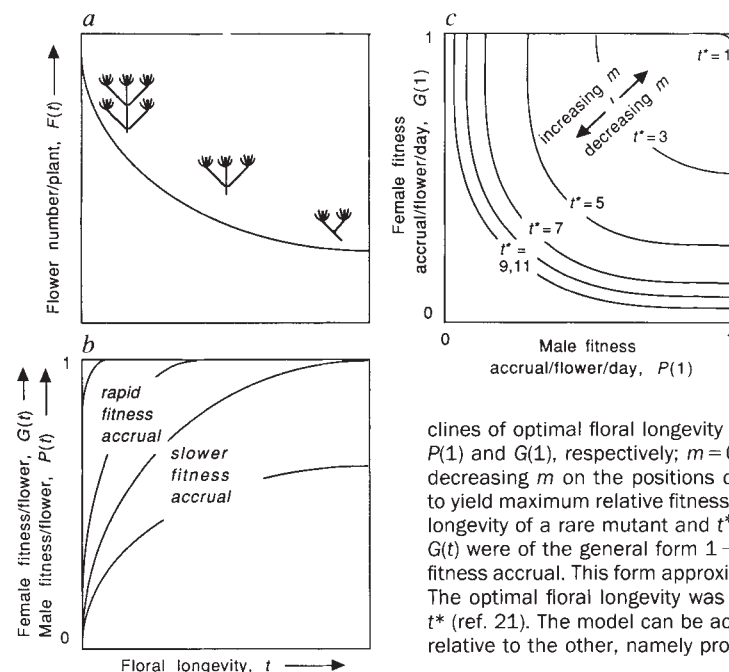
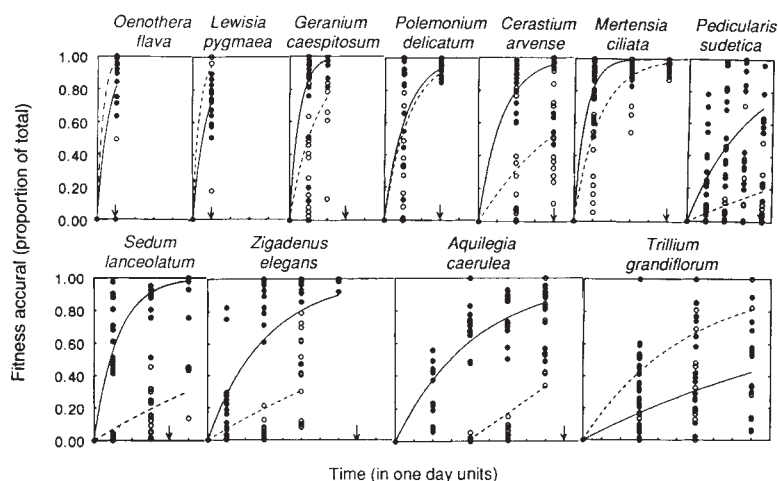


FIG. 1 Model for the evolution of optimal floral longevity. **a**, Flower number per plant, $F(t)$, declines with increasing floral longevity according to the function $F(t) = R/(c + tmc)$, where R is the total amount of resources available for flower construction, c is the amount of resources required for the construction of a single flower, m is the daily cost of maintaining a flower (for example, nectar production, respiratory maintenance) expressed as fraction of the cost of floral construction, and t is floral longevity (in days)⁴. **b**, Total male, $P(t)$, and total female, $G(t)$, fitness contributions of individual flowers as a function of floral longevity. $P(t)$ approaches 1 in a decelerating manner as all the pollen present in the flower is dispersed and added to the pool of pollen that competes for fertilization of ovules in the population. $G(t)$ likewise approaches 1 as the stigma receives sufficient pollen to ensure the fertilization of all the ovules in the flower. $P(t)$ and $G(t)$ may approach 1 rapidly or slowly, and may be strongly or weakly intercorrelated, depending for example on pollinator visitation rates and behaviour⁴. **c**, Isoclines of optimal floral longevity as a function of male and female fitness accrued per flower per day, $P(1)$ and $G(1)$, respectively; $m = 0.05$ for all isoclines, and arrows show the influence of increasing or decreasing m on the positions of the isoclines. Optimal floral longevity was found by solving for t to yield maximum relative fitness: $w = 1/2 \{F(t)/F(t^*)\} \{P(t)/P(t^*) + G(t)/G(t^*)\}$, where t denotes the floral longevity of a rare mutant and t^* denotes the evolutionarily stable strategy of floral longevity. $P(t)$ and $G(t)$ were of the general form $1 - e^{-rt}$, where r is a rate constant, and may differ for male and female fitness accrual. This form approximates well the fitness accrual data collected from natural populations. The optimal floral longevity was found by numerical solution of $dw/dt = 0$ and $d^2w/dt^2 < 0$, when $t = t^*$ (ref. 21). The model can be accommodated to incorporate delays in the start of one sexual function relative to the other, namely protandry and protogyny⁴.

FIG. 2 Male (filled circles, solid lines) and female (open circles, dashed lines) fitness accrual rates. Daily fitness accrual rates were determined by marking a random sample of flowers ($n = 22-84$) of each species at anthesis, and then harvesting a subset ($n = 6-10$) of these flowers at 1- or 2-day intervals over the flower lifetime. Each time interval was replicated twice. For flowers harvested, the number of pollen grains remaining in the anthers and adhering to the styles was determined. Random samples of flower buds from each species were also collected to determine the number of pollen grains initially present in the anthers. Raw data were transformed to reflect the accumulation of fitness on a proportional scale: that is, the proportion of the total pollen grains per flower present at $t=0$ that were disseminated after t days, and the proportion of the number of pollen grains required for full seed-set that were received after t days (as determined by the ovule number per flower multiplied by the average number of pollen grains required to fertilize an ovule (see ref. 3)). For each species, the pollen removal and deposition data were fitted by least-squares regression to functions of the form $1 - e^{-rt}$ using SYSTAT²². Arrows denote the end of a flower's lifetime



species were compared individually, the 95% confidence intervals of predicted and observed longevity typically did not overlap. Moreover, there is a consistent trend for predicted floral longevity to exceed those observed in nature. This may be due to underestimation of the cost of flower maintenance. In addition to carbon cost of flower maintenance and nectar production, as estimated in our study, there may be a maintenance cost associated with water loss due to flower transpiration and/or nectar production, and this could be substantial⁷. Additionally, investment in flower maintenance could reduce seed output if it draws upon resources necessary for seed maturation, or if it enhances the susceptibility of seeds to seed predators. The model can be revised to include these additional costs. Doing so would reduce

(senescence) as observed in the field (except *Trillium*, for which the lifetime was 21 days).

the predicted floral longevity and would improve the relationship between observed and predicted floral longevity, but because data on these and other maintenance costs are difficult to estimate, they are not included in our calculations of optimal floral longevity.

The model described and tested here assumes that plants adapt to low levels of pollinator activity through evolving differences in floral longevity. This is only one possible evolutionary outcome. For instance, alternative responses might include more or larger flowers per plant, which could act to increase pollinator visitation rates. But in cases where pollinators are rare to begin with, as in many orchids that rely on specific pollinators⁸, a 'sit-and-wait' strategy of increased floral longevity may be the only means for the flower to fulfil its reproductive role. As already noted, orchids often have long-lived flowers.

Flowers are only one of several plant organs that show variation in longevity. Variation in the longevity of fruits, leaves, shoots and roots also occurs and may be interpreted within a resource allocation and fitness accrual framework^{9,12}. For example, the photosynthetic ability of leaves changes during development; older leaves contribute less to whole plant carbon gain because of reduced photosynthetic ability and/or their position in the canopy (that is, self-shading). Other developed models suggest that there is an optimal longevity for leaves based on these parameters^{10,11}. There may be similar carbon-balance-driven causes of twig senescence¹². It may be generally fruitful to consider the optimal longevity of many plant structures within the framework of resource allocation, as we have shown here for flowers. □

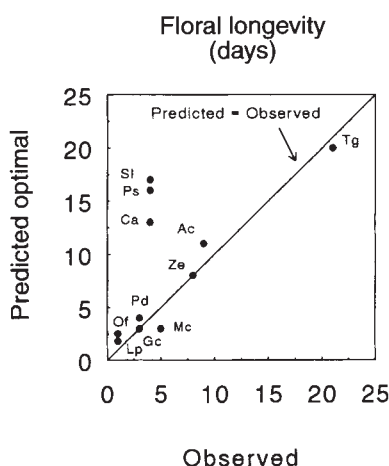


FIG. 3 Observed and predicted optimal floral longevity. Floral longevity was observed in 10–26 plants per species and rounded to the nearest day. Predicted floral longevity was obtained by substituting fitness accrual and cost parameters into the model as described in the text. Predicted floral longevity for *Aquilegia* was obtained from a form of the model modified to incorporate protandry⁴. The 95% confidence intervals for the observed and predicted floral longevity are as follows: *Aquilegia* (8.4–9.1 d, 10.8–13.5 d), *Cerastium* (4.2–4.7 d, 9.7–14.1 d), *Geranium* (1.1–1.5 d, 2.3–3.2 d), *Lewisia* (0.4–0.7 d, 1.5–2.2 d), *Mertensia* (4.6–5.0 d, 2.5–3.2 d), *Oenothera* (1.0–1.0 d, 1.9–3.8 d), *Pedicularis* (3.2–3.8 d, 11.5–23.5 d), *Polemonium* (2.2–2.9 d, 2.8–4.5 d), *Sedum* (4.0–4.6 d, 15.8–23.0 d), *Trillium* (20.1–21.9 d, 19.2–23.6 d), *Zigadenus* (7.1–8.1 d, 9.0–12.9 d). Confidence intervals for the predicted longevity were obtained by bootstrap sampling²³ of the data for fitness accrual rates (Fig. 2) and floral maintenance (Table 1), and using the values of m , $P(1)$, and $G(1)$ obtained in the solution for t^* . Repetition of this process generated a distribution of t^* .

Received 15 February; accepted 19 August 1994.

1. Kerner von Marilaun, A. *The Natural History of Plants, Their Forms, Growth and Distribution* (Holt, New York, 1895).
2. Primack, R. B. *A. Rev. Ecol. Syst.* **16**, 15–37 (1985).
3. Ashman, T.-L. & Schoen, D. J. in *Floral Biology* (eds Lloyd, D. G. & Barrett, S. C. H.) (Chapman and Hall, New York, in the press).
4. Schoen, D. J. & Ashman, T.-L. *Evolution* (in the press).
5. Charnov, E. L. *The Theory of Sex Allocation* (Princeton Univ. Press, Princeton, 1982).
6. Stearns, S. C. *The Evolution of Life Histories* (Oxford Univ. Press, Oxford, 1992).
7. Nobel, P. S. *Bot. Gaz.* **138**, 1–6 (1977).
8. Van der Pijl, L. & Dodson, C. H. *Orchid Flowers: Their Pollination and Evolution* (Univ. of Miami Press, Coral Gables, 1966).
9. Chabot, B. F. & Hicks, D. J. *A. Rev. Ecol. Syst.* **13**, 229–259 (1982).
10. Kikuzawa, K. *Am. Nat.* **138**, 1250–1263 (1991).
11. Ackerly, D. D. thesis, Harvard Univ. (1993).
12. Kasmer, J. M. *Bull. Ecol. Soc. Amer.* **74**, 302 (1993).
13. Miller, R. B. *Am. J. Bot.* **65**, 406–414 (1978).
14. Gerber, M. A. *Ecology* **66**, 762–772 (1985).
15. Williams, K., Field, C. B., & Mooney, H. A. *Am. Nat.* **133**, 198–211 (1989).
16. Chapin, F. S. *Am. Nat.* **133**, 1–19 (1989).
17. Lovett-Doust, J. & Harper, J. L. *New Phytol.* **85**, 251–264 (1980).
18. Ashman, T.-L. & Baker, I. *Ecology* **73**, 1237–1243 (1992).
19. Nesom, G. L. & La Duke, J. C. *Can. J. Bot.* **63**, 7–14 (1985).
20. Lubbers, A. E. & Lechowicz, M. J. *Ecology* **70**, 85–96 (1989).

21. Maynard Smith, J. *Evolution and the Theory of Games* (Cambridge Univ. Press, Cambridge, 1982).
22. Wilkinson, L. SYSTAT (SYSTAT Inc., Evanston, 1990).
23. Efron, B. *An Introduction to the Bootstrap*. (Chapman and Hall, New York, 1993).

ACKNOWLEDGEMENTS. We thank S. M. Barratt-Boyes, S. Dahl, C. Galen, M. Lechowicz and N. Wong for assistance; M. T. Morgan and M. L. Stanton for comments on the manuscript; and the Gault Estate Board of Directors for permission to conduct field research at Mont St Hilaire. This work was supported by an NSERC operating grant to D.J.S., an NSERC International Postdoctoral Fellowship to T.L.A., and a grant from Fonds pour la Formation de Chercheurs et l'Aide à la Recherche (Québec). Additional logistic support was provided by the Department of Biological Sciences, University of Pittsburgh.

Parallel detection of Kanizsa subjective figures in the human visual system

Greg Davis & Jon Driver

Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, UK

SUBJECTIVE figures, seen in the absence of luminance gradients (Fig. 1), provide a phenomenal illusion that can be related to the properties of single cells in the visual cortex¹, offering a rare bridge between brain function and visual awareness. It remains controversial whether subjective figures arise from intelligent cognitive mechanisms²⁻⁴, or from lower-level processes in early vision^{5,6}. The cognitive account implies that the perception of subjective figures may require serial attentive processing, whereas on the low-level account they should arise in parallel at earlier visual stages. Physiological evidence apparently fits a low-level account^{7,8} and indicates that some types of subjective contour may be detected earlier than the conventional Kanizsa⁹ type. Here we report that even Kanizsa subjective figures can be detected without focal attention at parallel stages of the human visual system.

Two previous studies have examined human visual search for subjective figures to determine whether they are coded in parallel or by a serial attentional process applied to each element in turn^{10,11}. Parallel search has been reported for subjective contours defined by offset gratings (Fig. 1a), suggesting coding at an early stage without focal attention¹⁰. Consistent with this low-level account, subjective contours of this type elicit responses from macaque neurons in V1⁷, which is the first stage of cortical visual processing. However, there are several potential difficulties in interpreting results from offset-grating stimuli alone (Fig. 1a legend).

In contrast to the finding for offset gratings, human visual search for Kanizsa figures (Fig. 1b) has been reported to be serial¹¹, implying that attention is required in their construction. Thus, different mechanisms may be responsible for the different types of subjective figure. The single-cell results appear to be consistent with this; unlike offset-grating stimuli, neural responses to Kanizsa figures have not been found as early as primate V1, arising instead in V2⁸ (Fig. 1c, d).

These V2 responses occur even when the monkey's attention is engaged in a visual task at a different location, however^{1,8}. This suggests that Kanizsa contours can be coded without focal attention in primates, in apparent conflict with the serial visual search¹¹ in humans. Accordingly, we re-examined whether human visual search for Kanizsa figures can be parallel when previous methods are refined to yield optimal conditions.

Any tendency to group together inducers from distinct figures (as a result of their proximity or common colour) could obscure a capacity for parallel coding, because visual search is dramatically impaired when elements of the target are grouped with distractor elements¹². To circumvent this problem, we segregated the display into distinct candidate figures in advance, by the use of circular place-holders (Fig. 2a). In addition to an appropriate

grouping of inducers, the perception of Kanizsa subjective figures may also require a figure-ground reversal¹⁴, such that the inducing 'pac-men' become the background for the subjective figure. In the previous search study with Kanizsa stimuli¹¹, the inducers appeared abruptly on a blank screen, and so might initially be parsed as figures. The need to reverse this assignment may have been responsible for the serial processing observed. In our task, the inducers are initially present as uninformative filled circles (Fig. 2a) and segments are then abruptly removed to yield inducing pac-men (Fig. 2b). As a result, the informative segment is figural from its onset. Finally, as the perception of Kanizsa figures falls off sharply with retinal eccentricity¹¹, we equated this factor across search items by presenting all our stimuli on an imaginary horizontal ellipse centred at fixation (Fig. 2).

With these departures from previous work¹¹, experiment 1 examined visual search for a Kanizsa square among non-targets that were made up from the same inducers but rearranged to produce no subjective figure (Fig. 2b). The results (see E1 in Fig. 3) show that average reaction times were scarcely affected by the number of non-targets, falling within the range conventionally taken to indicate parallel processing¹³⁻¹⁶.

A minor alteration to these stimuli (narrow arcs across the inducing segments; see E2 in Fig. 3) removes the percept of a subjective square, and also blocks the V2 response to Kanizsa contours^{1,8} (Fig. 1d). Experiment 2 shows that this minimal alteration led to strongly serial search (see graphs for E2 in Fig. 3), suggesting that subjective figures were indeed responsible for the parallel results in experiment 1.

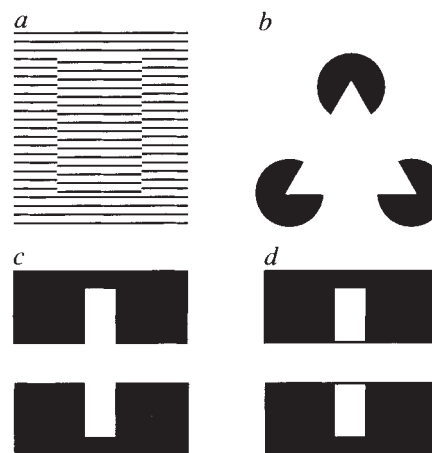


FIG. 1 Examples of stimuli that induce the perception of a subjective figure, including contours that are not defined by luminance gradients. *a*, Abutting offset gratings, with two vertical subjective contours. Variants of such contours have been shown to elicit responses in single cells within area V1 in macaque visual cortex⁷, and to yield parallel visual search in humans¹⁰. Both results suggest coding at early stages of vision. Unfortunately, there are some interpretative problems with offset-grating stimuli. First, the subjective contours within them may be atypical, because they could be signalled by the alignment of multiple bar-terminations, which are coded at the earliest levels of the visual system. Second, the offset gratings met at an oblique angle in the human visual search study¹⁰, thus producing low spatial-frequency artefacts at their intersection. *b*, Conventional Kanizsa-style subjective figure, named after the Italian psychologist who introduced them⁹. Subjective contours produced by such inducers may have a different neural origin^{7,8} from those in *a*, and have been reported to yield serial search¹¹. *c*, Simplified Kanizsa stimulus that has been shown to elicit responses in neurons whose entire receptive field falls inside the 'gap' between inducers (that is, the responses are to the vertical subjective contours that bridge this gap, not to the inducers themselves). This has been found for cells in macaque V2, but not for V1^{7,8}. *d*, Control stimulus; the additional thin lines that close the inducers eliminate the subjective percept, and also eliminate the macaque V2 response^{1,8}. A variant on this control is used in experiment 2.

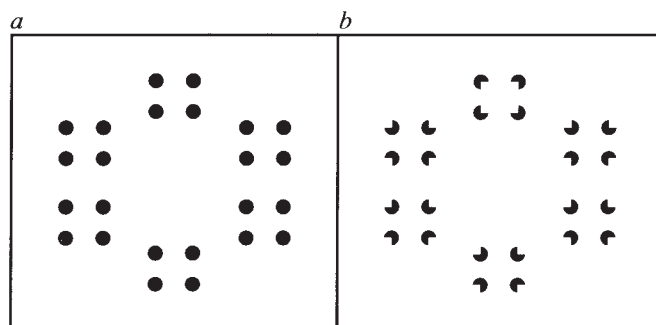
Those parallel results might reflect 'blurring' together of the black target inducers (around the white region they circumscribe) within the low spatial-frequency channels of the primate visual system, however¹⁷. This explanation seems unlikely, given the dramatic effect of the thin arcs in experiment 2, but we nevertheless ran two further studies to rule it out. Linear low spatial-frequency filters should not blur inducers together when they have opposite contrast-polarity, but such inducers nevertheless yield subjective figures with unrestricted viewing (albeit with decreased subjective brightness¹⁸). Using opposite-polarity inducers (Fig. 4), experiment 3 replicated the parallel search for Kanizsa squares (E3 in Fig. 3).

Unfortunately, earlier work has suggested that human spatial frequency channels may actually combine signals with opposite

contrast in a nonlinear fashion¹⁹. This might in principle allow a low spatial-frequency channel to blur inducers together, even when they have opposite contrast²⁰ as in experiment 3. Accordingly, our final study used a different method to eliminate the low spatial-frequency account. Opposite-polarity inducers were again employed, but the target was now an outline subjective square (E4 in Fig. 3). Note that the narrow inducing segments in this case would necessarily be blurred out in low spatial-frequency channels that could bridge across the inducers. Nevertheless, parallel search was observed once again.

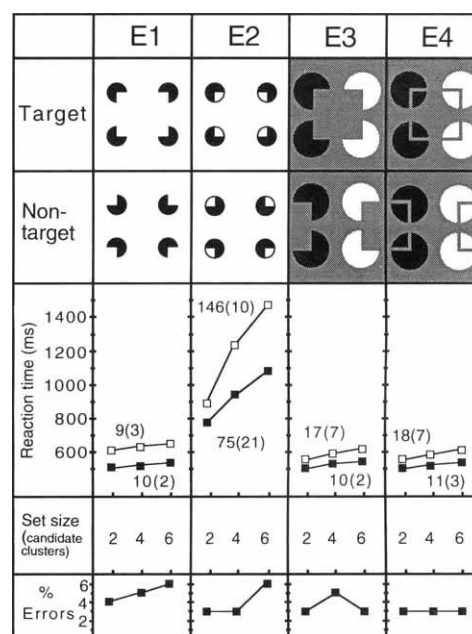
These findings demonstrate that human perception of Kanizsa subjective figures does not require focal attention. They support theories on which subjective figures derive from parallel, automatic processes at early stages of the visual system^{5,6}. Subjective

FIG. 2 Examples of the two successive displays within each trial of experiment 1, both to scale. *a*, The place-holder display, which began every trial following a 255-ms central fixation cross. These displays always contained 6 clusters of 4 solid black circles each, arranged as shown. Each circle was 0.4° in diameter at the viewing distance of 112 cm. *b*, The search display, which immediately replaced the place-holder display 900 ms after its onset, and remained until response. The search display was created by removing a segment from each place-holder circle within 2, 4 or 6 clusters (this is illustrated for 6 clusters, but the three set-sizes were equally likely) so that these clusters became candidates for the search process. At the smaller set sizes, the candidate clusters were diametrically opposite. On half the trials, a single subjective square target was present (illustrated at the top in *b*, but equally likely in the 6 cluster-locations) among candidate clusters made up of the same inducers but with a different spatial arrangement that yielded no subjective figure. For the remaining target-absent trials, all the candidate clusters had the latter arrangement. Subjects specified the presence or absence of the target as quickly and accurately as possible, via the index or middle finger of their preferred hand respectively. All experiments were performed on a Macintosh Quadra 610 with Trinitron colour monitor using Vscope software²². The monitor was set



to minimum brightness and maximum contrast, conditions that favour the perception of clear subjective figures under unrestricted viewing²³. Each subject underwent 16 blocks of 60 trials with target presence, target position, and the number of candidate clusters in randomized sequence. The first four blocks were excluded as practice.

FIG. 3 There were 12, 8, 7 and 6 different subjects in experiments 1, 2, 3 and 4 respectively. Each column depicts stimulus elements and results for a single experiment. A target is shown at the top, and then a nontarget, both to scale. In E3 and E4 these were equally likely to have the opposite contrast-polarity to that shown (across trials for targets, within trials for nontargets). The upper graph plots mean reaction time against the number of candidate clusters, separately for correct target-present responses (filled symbols) and correct target-absent responses (open symbols), with the slope of the best-fitting straight line (in ms per candidate cluster) given numerically in each case to indicate search rate, followed in parentheses by the standard error of this slope across the group of subjects. The error slopes did not differ significantly from zero for experiments 1, 3 and 4. Whereas the *RT* slopes did differ slightly from zero, this is expected mathematically^{15,16} for any parallel process which is subject to a degree of random noise (as in the case of the nervous system). Shallow target-present slopes of around 10 msec/item are conventionally taken to indicate parallel search^{13,14}. Experiments 1, 3 and 4 fall in this parallel range. Note that the quoted slopes (those analysed) are conservative in this respect, because they consider clusters as a single unit; slopes calculated against the number of inducers instead would yield 1/4 of the cited values. The slopes for experiment 2 were significantly steeper than experiments 1, 3 and 4 ($F(1, 18) = 62.0$, $F(1, 13) = 35.0$, and $F(1, 12) = 30.0$ respectively for 'Yes' slopes in planned contrasts, all values of $P < 0.0001$; and $F(1, 18) = 49.7$, $F(1, 13) = 16.8$, $F(1, 12) = 20.5$ respectively for 'No' slopes, all values of $P < 0.002$). Only the slopes for experiment 2 fall within the range conventionally taken to indicate serial search^{13,14}, and have the further serial, self-terminating hallmark of a 2:1 ratio for target-absent versus target-present slopes^{13,15}. Indeed, the target-absent and target-present slopes for this experiment differed significantly from each other ($F(1, 7) = 22.5$, $P = 0.002$) but not from a 2:1 ratio ($F(1, 7) < 1$, NS). By contrast, the target-present and target-absent slopes did not differ significantly (all values of $F < 2.5$, NS) within the other three studies. The apparently near 2:1 ratio for target-absence versus target-presence in the mean slopes for experiments 3 and 4 were in both



cases due to a single subject with a steep target-absent slope (presumably caused by a serial checking strategy when no target was detected¹³) which inflated the mean. Excluding these single subjects from experiments 3 and 4 yields target-present slopes of 8 ms per cluster and 9 ms per cluster respectively, and target-absent slopes of 11 ms per cluster and 9 ms per cluster respectively.

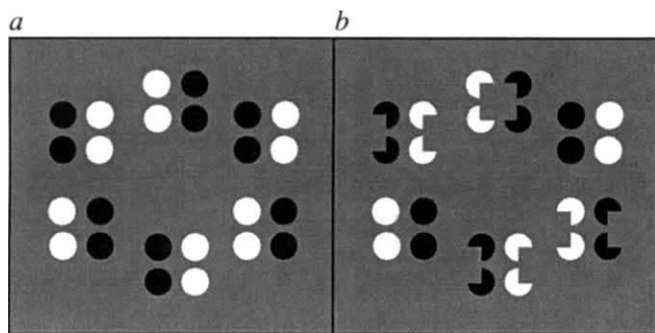


FIG. 4 Examples of the two successive displays within each trial of experiment 3, both to the same scale as in Fig. 2. *a*, The place-holder display, similar to experiments 1 and 2 except for the use of opposite-contrast circles against an intermediate grey background. Note that the side of black and of white circles is counterbalanced across clusters. Moreover, the arrangement of black and white circles was equally likely to be the opposite of that shown, and this was unpredictable across trials. *b*, The search display, which as before immediately replaced the place-holder display on its termination, and which again corresponded to the removal of 4 segments from each of a varied number of clusters (a set size of four candidate clusters is illustrated, but two or six were equally likely). Two of the inducers within the square target (shown at top) have common polarity. Note, however, that their arrangement is now shared with the nontarget clusters, so that the target is specified only by its arrangement of opposite-polarity inducers. Moreover, the side of black and white inducers for each cluster was both unpredictable across trials, and heterogeneous within trials, precluding any strategy of searching for a common-polarity subjective contour in a particular location. In all other respects, the method followed experiments 1 and 2. Experiment 4 was identical to experiment 3 except that different (substantially smaller) segments were removed from the place-holder displays to yield a search display that might contain an outline square target.

figure perception is thought to provide an anticamouflage device²¹ which detects the non-accidental properties of otherwise hidden objects. Our parallel results make functional sense from this perspective, as such a device would be self-defeating if it required serial attentional scrutiny to detect camouflaged objects. Our findings do not, however, preclude a role for higher, possibly attentional, factors^{2,4}. These may modulate the output of the parallel stages identified here so that, for instance, subjective figures are not seen when occlusion relations⁴, or the completeness of familiar inducers⁹, argue against the existence of an object under camouflage. □

Received 12 May; accepted 5 September 1994.

- Peterhans, E. & von der Heydt, R. *Trends Neurosci.* **14**, 112–119 (1991).
- Gregory, R. L. *Nature* **238**, 51–52 (1972).
- Coren, S. *Psych. Rev.* **79**, 359–367 (1972).
- Rock, I. & Anson, R. *Perception* **8**, 665–681 (1979).
- Brigner, W. L. & Gallagher, M. B. *Percept. Motor Skills* **38**, 1047–1053 (1974).
- Smith, A. T. & Over, R. *Nature* **257**, 581–582 (1977).
- Grosf, D. H., Shapley, R. M. & Hawken, M. J. *Nature* **365**, 550–552 (1993).
- von der Heydt, R., Peterhans, E. & Baumgartner, G. *Science* **224**, 1260–1262 (1984).
- Kanizsa, G. *Rivista di Psicologia* **49**, 7–30 (1955).
- Gurnsey, R., Humphrey, G. K. & Kapitan, P. *Percept. Psychophys.* **52**, 263–276 (1992).
- Grabowecky, M. & Treisman, A. *Invest. Ophthalm. Vis. Sci. suppl.* **30**, 457 (1989).
- Duncan, J. & Humphreys, G. W. *Psych. Rev.* **96**, 433–458 (1989).
- Treisman, A. M. & Gelade, G. *Cog. Psych.* **12**, 97–136 (1980).
- Enns, J. T. & Rensink, R. A. *Science* **247**, 721–723 (1990).
- Townsend, J. T. *Percept. Psychophys.* **10**, 161–163 (1971).
- Broadbent, D. E. in *Modelling Cognition* (ed. Morris, P.) 169–186 (Wiley, London, 1985).
- Ginsburg, A. P. *Nature* **257**, 219–220 (1975).
- Pradny, K. *Percept. Psychophys.* **34**, 403–404 (1983).
- De Valois, R. L. & De Valois, K. K. *Spatial Vision* (Oxford Univ. Press, Oxford, 1990).
- Ginsburg, A. P. in *The Perception of Illusory Contours* (eds Petry, S. & Meyer, G. E.) 126–130 (Springer, New York, 1987).
- Ramachandran, V. L. in *The Perception of Illusory Contours* (eds Petry, S. & Meyer, G. E.) 93–108 (Springer, New York, 1987).
- Enns, J. T., Ochs, E. P. & Rensink, R. A. *Behavior Research Instruments & Computers* **22**, 118–122 (1990).
- Warm, J. S., Dember, W. N., Padich, R. A., Beckner, J. & Jones, S. *Percept. Psychophys.* **39**, 219 (1986).

ACKNOWLEDGEMENTS. We thank C. Rorden and M. East. This work was supported by a grant from the Science and Engineering Research Council (UK).

Broad tuning for spatial frequency of neural mechanisms underlying visual perception of coherent motion

Yuede Yang & Randolph Blake

Department of Psychology, Vanderbilt University, Nashville, Tennessee 37240, USA

NEURAL events underlying perception of coherent motion are generally believed to be hierarchical^{1,2}: information about local motion is registered by spatio-temporal coincidence detectors^{3–5} whose outputs are cooperatively integrated at a subsequent stage^{6,7}. There is disagreement, however, concerning the spatial scale of the neural filters underlying these operations. According to one class of models, motion registration is initially accomplished in parallel at multiple spatial scales^{3–5}, with filters tuned to lower spatial frequencies responsive to larger motion displacements than filters tuned to higher frequencies. According to another scheme, motion analysis involves a single, broadly tuned spatial filter, with optimal displacement dependent on spacing of local elements⁸. Here we use a masking procedure to measure the extent to which dynamic noise depicted at one spatial scale interferes with detection of coherent motion conveyed by image features at another spatial scale. Our results indicate that a single filter, broadly tuned for spatial frequency, is mediating detection of coherent motion. This finding dovetails with known physiological properties of neurons at an intermediate stage of motion processing.

Observers viewed two successive animation sequences each depicting a circular array of dots which were displaced from frame-to-frame of the sequence. In one sequence all dots were displaced upwards, yielding the perception of smooth, coherent motion, whereas in the other they were displaced in random directions, yielding incoherent motion; observers simply indicated which interval portrayed coherent motion, guessing if necessary. Superimposed on these two types of animation sequences (coherent and random motion) were dynamic noise dots that could be made sufficiently strong to destroy (mask) the perception of coherent motion, thus rendering the discrimination task impossible. Using spatial-frequency filtering techniques⁹, the successive frames of the coherent motion sequence, the random motion sequence and the noise masks were band-pass filtered so that their spatial frequency content was only 0.4 octaves wide; the centre frequency of the noise was set independently of the centre frequency of the coherent and the random motion (Fig. 1). The root-mean-square (r.m.s.) contrast of the coherent and random motion sequences remained constant at 0.35; r.m.s. contrast of the mask was varied to find the value yielding 75% correct performance on this two-alternative, forced-choice task. In our initial experiment, dot displacement for the coherent motion sequence was approximately 1/2 cycle of the centre spatial frequency of the filtered sequence under test¹⁰.

Viewed on their own without the addition of noise, these filtered animation sequences could be easily discriminated from uncorrelated sequences. Adding noise to this strong motion signal degraded the ability to detect coherent motion (Fig. 2), even when the spatial frequency content of the noise differed greatly from that of the coherent motion sequences. Applying statistical procedures⁹, we found that all resulting masking tuning functions peaked at approximately the same value of 4 c deg⁻¹ and that all had about the same full-width tuning of 2.4 octaves at the 3 dB roll-off point. These results implicate a single neural filter underlying coherent motion perception, a filter broadly tuned for spatial frequency. This conclusion is consistent with the single-channel model proposed by Morgan⁸. It is arguable, however, whether the filter implicated in our work operates at

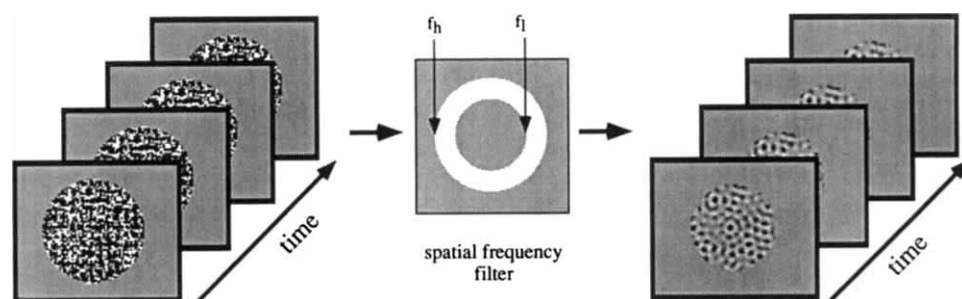


FIG. 1 Random-dot cinematograms (RDCs) were generated on a 19-inch radius grey-scale monitor (1152H $\times$ 882V pixel resolution; P104 phosphor) under control of a Macintosh II computer. Calibrated look-up tables were corrected for luminance nonlinearity. Each original unfiltered RDC consisted of ten successive frames (each of 84-ms duration), with each frame containing a circular area filled with 50% black (0.17 cdm^{-2}) and 50% white (74 cdm^{-2}) dots; the surrounding screen region was uniform grey (26 cdm^{-2}). Each dot was 2×2 pixels which, at the 0.8-m viewing distance, subtended $2.66 \text{ min arc} \times 2.66 \text{ min arc}$. The diameter of the circular, texture-filled region was 2.7 deg arc . Coherent motion was created by displacing all dots a specified number of pixels upwards. A schematic example of several frames of this

complex (unfiltered) RDC is shown on the left. RDCs consisting entirely of random motion (no net movement in any given direction) were also created and constituted the incoherent motion and the noise masks. RDCs composed of narrow bands of spatial frequencies were created by filtering each of the ten frames of an RDC with a 0.4-octave isotropic filter (middle, with f_l and f_h referring to the low- and high-frequency cut-off). Filter size was 256×256 pixels, with highest spatial frequency 22 c deg^{-1} at the viewing distance of 0.8 m. A schematic example of several frames of the resulting RDCs are shown on the right. Temporal frequency remained constant (and speed varied) because frame-to-frame displacement varied inversely with centre spatial frequency. Incoherent motion and noise masks were comparably filtered.

the initial stage of motion registration as Morgan proposed, or at a subsequent stage at which local motion signals are integrated. It is noteworthy that the single, broadly tuned filter implicated in our work could register coherent motion signalled exclusively by high spatial frequencies, for others have shown that high spatial frequencies contribute to motion perception^{11,12}.

It could be argued that the 1/2-cycle displacement used in our animation sequences was not optimal for idealized motion-energy detectors^{3,5}, which are maximally responsive to 1/4-cycle displacements. We therefore repeated the masking experiment using dot displacements that were always 1/4 cycle of the centre spatial frequency of the sequence. Testing at three different centre spatial frequencies, we found that results were essentially

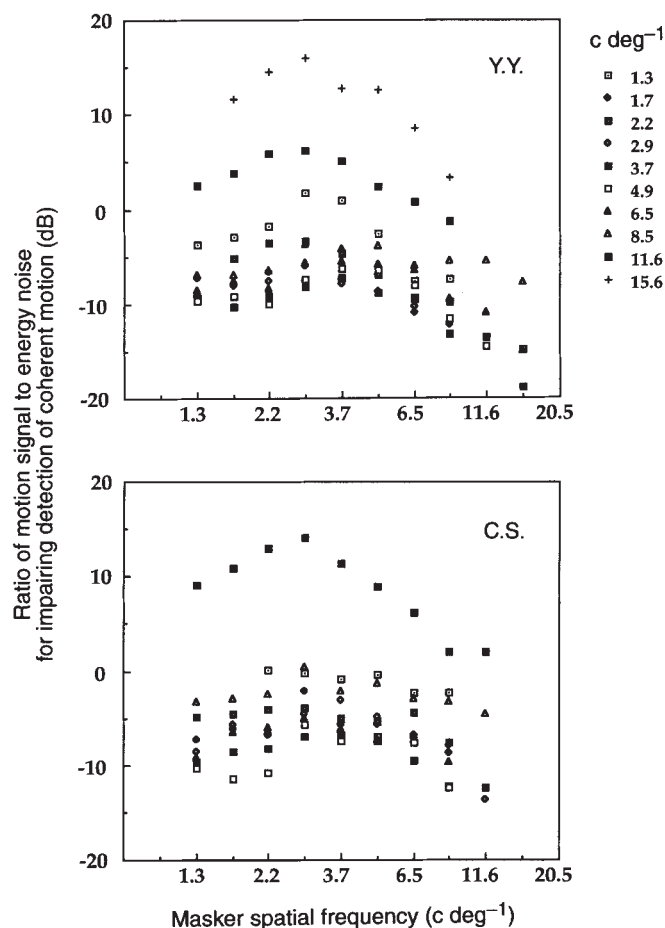


FIG. 2 Results from masking experiment. On each trial the observer viewed two successive animation sequences (each lasting 0.84 s and separated by a 1-s interval), one sequence being coherent upward motion and the other being incoherent random motion. The dots comprising both animation sequences were narrowly filtered with the same centre spatial frequency and bandwidth. The r.m.s. contrast of both types of animation sequence was held constant at 0.35. Following each trial, the observer judged which interval (first or second) contained the coherent motion sequence, with feedback. Narrowly filtered masking noise was added to both animation sequences, and the r.m.s. energy level of this masker varied from trial-to-trial within a range yielding seven signal-to-noise values (calculated as $20 \times \log (\text{r.m.s.}_{\text{signal}} / \text{r.m.s.}_{\text{masker}})$), each differing by 1 dB. Over blocks of trials, the centre spatial frequency of the masker was varied relative to the centre spatial frequency of the motion sequences. Probit analysis was used to fit cumulative gaussian curves to the resulting psychometric functions, and the signal-to-noise value associated with the 75% point defined the masked threshold with this 2-interval forced-choice procedure. The graphs plot these threshold signal-to-noise values as a function of masker centre spatial frequency for two observers (one author and one naive). Each set of data points corresponds to a given centre spatial frequency of coherent motion, as denoted by the symbols shown to the right of the upper graph. Large signal-to-noise values indicate that relatively little noise energy noise was required to mask coherent motion. For these results, dot displacement was 1/2-cycle of the centre spatial frequency of the filtered coherent motion signal¹⁰. In a second experiment, essentially identical results were obtained using a 1/4-cycle displacement.

the same as those shown in Fig. 2, implying that the broadness of tuning is not attributable to the particular displacement used. The spatial frequency selectivity underlying the registration of coherent motion differs from that found for stereopsis⁹, for detection of drifting gratings¹³ and for integration of drifting cosine gratings into coherent plaids¹⁴.

A third, ancillary experiment measured sensitivity to coherent motion by varying the percentage of dots moving coherently upwards; again the coherent motion sequences and the random-motion sequences were both narrowly filtered, but masking noise was not added to either sequence. Sensitivity to coherent motion varies with spatial frequency (Fig. 3), being highest near 4 c deg^{-1} , the same as the peak value implicated in our masking experiment.

Implications of our results are usefully portrayed in a space/time format¹⁵. For unfiltered animation sequences, upward coherent motion appears as diagonally oriented contours in the space/time domain (Fig. 4A). The two-dimensional Fourier power spectrum of this stimulus is defined by a streak of spatial and temporal frequency energy (Fig. 4B), whose orientation in the spectrum is proportional to motion speed. For coherent motion sequences that have been spatial frequency filtered, this streak is reduced to two small islands of spatio-temporal energy within the two-dimensional Fourier spectrum (Fig. 4C). The unfiltered noise masker, in comparison, consists of a disorganized, isotropic 'cloud' of dots in space/time (Fig. 4D) whose two-dimensional Fourier spectrum is likewise isotropic (Fig. 4E). When filtered, energy in the two-dimensional spectrum of the noise is broadly distributed in temporal frequency but narrowly distributed in spatial frequency (Fig. 4F). Presumably, masking (impaired perception of coherent motion by addition of dynamic noise) involves obliteration of the 'island' of spatio-

temporal energy associated with coherent motion (Fig. 4C). Our results show that this destructive interference occurs even when signal and noise occupy different regions of the two-dimensional energy spectra. In other words, the filters underlying detection of coherent motion have relatively broad spectral bandwidths. This makes computational sense, for broad spatial frequency tuning implements integration of local motion signals over space, a requisite for perception of coherent motion^{11,16}. Exactly where these broadly tuned filters fit into the emerging picture of multiple motion pathways^{17,18} remains to be determined. As for the actual neural bases of the mechanism isolated and studied here, the broad spatial frequency tuning for coherent motion is a property

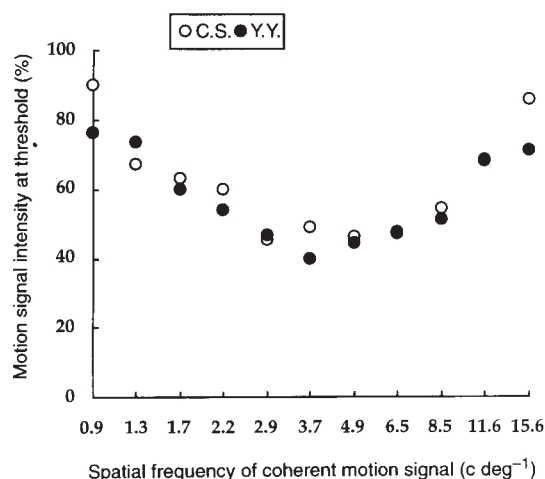


FIG. 3 To assess detectability of coherent motion signalled at different spatial scales, the percentage of RDC dots moving upwards was varied to determine the minimum signal strength supporting 75%-correct detection of coherent motion of filtered (but unmasked) RDCs. On each trial in a 2-interval forced-choice experiment, one interval displayed uncorrelated motion (0% coherent motion signal) and the other displayed coherent motion whose signal strength (percentage of upward-moving dots) varied over trials; the observer indicated which interval contained coherent motion, with feedback given. Signal dots were displayed by an amount equal to $\sim 1/2$ -cycle of the centre spatial frequency¹⁰. Resulting thresholds (determined by probit analysis) for two observers reveal that sensitivity for coherent motion is greatest near 4 c deg^{-1} . Coherence thresholds here are somewhat higher than those reported previously for unfiltered RDCs¹⁶, probably because the much higher dot density in our animation sequences creates more potential false matches.

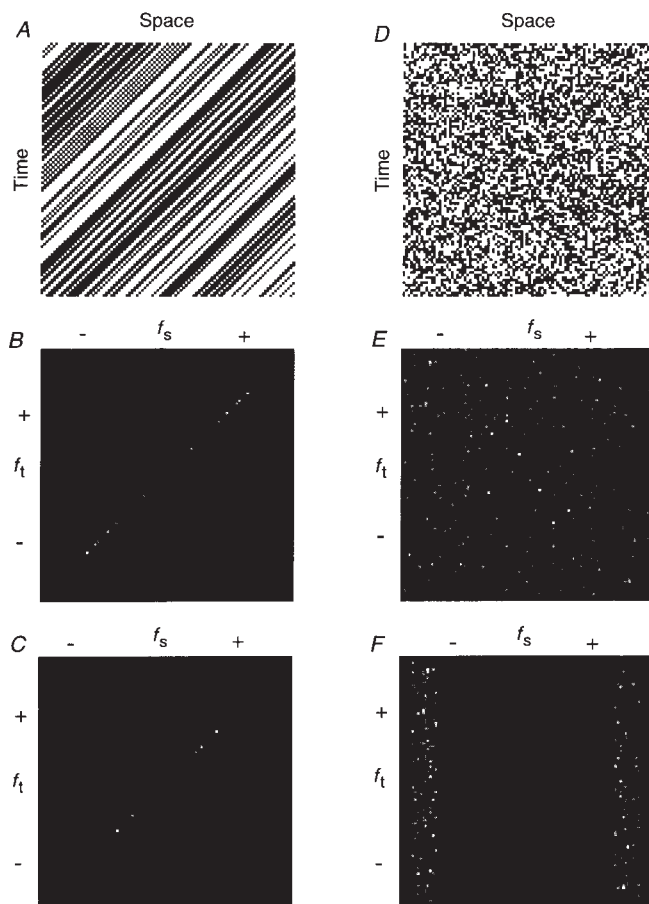


FIG. 4 Space/time portrayals of coherent motion and masking noise, together with their two-dimensional (2-D) Fourier power spectra. A, Plot showing successive vertical positions of unfiltered dots (50% black/50% white) moving coherently upwards in fixed displacements from frame to frame at a constant speed. In this plot the bottom row of dots represents dot positions at time t_1 and successively higher rows denote the vertical positions of those dots at times $t_2 \dots t_n$. Because all dots move coherently upward, any given dot's position does not change horizontally over time. B, 2-D Fourier power spectra of the event depicted in A. Because dots were unfiltered, energy is distributed broadly in spatial frequency and in temporal frequency. C, 2-D Fourier power spectra of coherent motion of spatially filtered dots. Energy is concentrated in small 'islands' of the spectra, corresponding to the centre frequency of the bandpass of the filter. Limiting spatial frequency content also limits temporal frequency. D, Space-time plot for dynamic masking noise, where individual dots are simply replotted from frame-to-frame; perceptually, this event looks like random snow on a detuned television. E, 2-D Fourier power spectra of the unfiltered masking noise. F, 2-D Fourier power spectra of spatially filtered masking noise. Energy is concentrated in spatial frequency but widely distributed in temporal frequency.

more characteristic of neurons in the middle temporal visual area, not neurons in area VI¹⁹. □

Received 15 July; accepted 2 September 1994.

1. Movshon, A. in *Images and Understanding* (eds Barlow, H., Blakemore, C. & Weston-Smith, M.) (Cambridge University Press, Cambridge, 1990).
2. Braddick, O. *Trends Neurosci.* **16**, 263–268 (1993).
3. Adelson, E. H. & Bergen, J. R. *J. opt. Soc. Am.* **A2**, 284–299 (1985).
4. Van Santen, J. P. H. & Sperling, G. *J. opt. Soc. Am.* **A1**, 451–473 (1984).
5. Watson, A. B. *et al. J. opt. Soc. Am.* **A3**, 300–307 (1986).
6. Chang, J. J. & Julesz, B. *Vision Res.* **24**, 1781–1788 (1984).
7. Nawrot, M. & Sekuler, R. *Vision Res.* **30**, 1439–1451 (1990).
8. Morgan, M. J. *Nature* **355**, 344–346 (1992).
9. Yang, Y. & Blake, R. *Vision Res.* **31**, 1177–1190 (1991).
10. Cleary, R. & Braddick, O. *Vision Res.* **30**, 303–316 (1990).
11. Smith, A. T. *et al. Vision Res.* **34**, 2425–2430 (1994).
12. van de Grind, W. A. *et al. Expl. Brain Res.* **91**, 135–150 (1992).
13. Anderson, S. J. & Burr, D. C. *Vision Res.* **13**, 1147–1154 (1985).
14. Adelson, E. H. & Movshon, J. A. *Nature* **300**, 523–525 (1982).
15. Adelson, E. H. *Opt. Photon. News* **2**, 24–30 (1991).
16. Williams, D. W. & Sekuler, R. *Vision Res.* **24**, 55–62 (1984).
17. Werkhoven, P., Sperling, G. & Chubb, C. *Vision Res.* **33**, 463–485 (1993).
18. Wilson, H. R. & Kim, J. *Vision Res.* **34**, 1835–1842 (1994).
19. Britten, K. H. *et al. Vis. Neurosci.* **10**, 1157–1169 (1993).

ACKNOWLEDGEMENTS. Supported by a research grant (R.B.) and a Vision Core grant from the NIH.

Krox-20 controls myelination in the peripheral nervous system

Piotr Topilko*, Sylvie Schneider-Maunoury*, Giovanni Levi*†, Anne Baron-Van Evercooren‡, Amina Ben Younes Chennoufi‡, Tania Seitanidou*, Charles Babinet§ & Patrick Charnay*

*Unité 368 de l'Institut National de la Santé et de la Recherche Médicale, Ecole Normale Supérieure, 46 rue d'Ulm, 75230 Paris Cedex 05, France

†Advanced Biotechnology Center, Viale Benedetto XV n°10, 16132 Genova, Italy

‡Unité 134 de l'Institut National de la Santé et de la Recherche Médicale, Hôpital de la Salpêtrière, 75651 Paris Cedex 13, France

§Unité de Recherche Associée 361 du Centre National de la Recherche Scientifique, Institut Pasteur, 75724 Paris Cedex 15, France

THE molecular mechanisms controlling the process of myelination by Schwann cells remain elusive, despite recent progress in the identification and characterization of genes encoding myelin components (reviewed in ref. 1). We have created a null allele in the mouse *Krox-20* gene, which encodes a zinc-finger transcription factor^{2,3}, by in-frame insertion of the *Escherichia coli lacZ* gene, and have shown that hindbrain segmentation is affected in *Krox-20*^{-/-} embryos⁴. We demonstrate here that *Krox-20* is also activated in Schwann cells before the onset of myelination and that its disruption blocks Schwann cells at an early stage in their differentiation, thus preventing myelination in the peripheral nervous system. In *Krox-20*^{-/-} mice, Schwann cells wrap their cytoplasmic processes only one and a half turns around the axon, and although they express the early myelin marker, myelin-associated glycoprotein⁵, late myelin gene products are absent, including those for protein zero⁶ and myelin basic protein⁷. Therefore *Krox-20* is likely to control a set of genes required for completion of myelination in the peripheral nervous system.

In late embryos, the *Krox-20/lacZ* chimaeric gene is expressed in a complex pattern revealing a number of potential sites of *Krox-20* expression (P.T. *et al.*, manuscript in preparation), including part of the peripheral nervous system (PNS). Between 10 and 15 days after birth, surviving *Krox-20*^{-/-} mutant mice

tremble⁴, suggesting a possible phenotypic effect of the mutation on the PNS, so we analysed *Krox-20/lacZ* expression in the PNS in greater detail. No qualitative differences in β -galactosidase activity were observed between heterozygotes and homozygotes during embryogenesis, although the intensity of the labelling was stronger in homozygotes (data not shown). In the PNS, β -galactosidase is first detected in boundary cap cells that surround nerve exit points from the central nervous system (CNS) around 10.5 days post-coitum (d.p.c.)^{4,8}. Up to 14.5 d.p.c., PNS expression is confined to the motor and sensory roots of cranial and spinal nerves, with no labelling of the ganglia or of more distal parts of the fibres (Fig. 1a), with the exception of the Xlth cranial nerve which is labelled along its path adjacent to the CNS⁴. A major transition occurs around 15.5 d.p.c., when *Krox-20/lacZ* is activated along the entire length of peripheral nerves (Fig. 1b, c). However, no β -galactosidase is detected within cranial and dorsal root ganglia or in the sympathetic system (Fig. 1 and data not shown). This expression persists during the entire life of the animal. Myelination in the PNS occurs during the first 2 weeks after birth, and in heterozygous animals, β -galactosidase activity is detected in the cytoplasm of myelinating Schwann cells (Fig. 1d). *In situ* hybridizations on sections from wild-type animals confirmed that *Krox-20/lacZ* expression in peripheral nerves reflects normal expression of *Krox-20* (Fig. 1e). This distribution strongly suggests that expression in the PNS is restricted to myelinating Schwann cells and putative Schwann cell precursors⁹ and, indeed, purified Schwann cells from newborn sciatic nerves express *Krox-20* (P. Murphy *et al.*, manuscript in preparation).

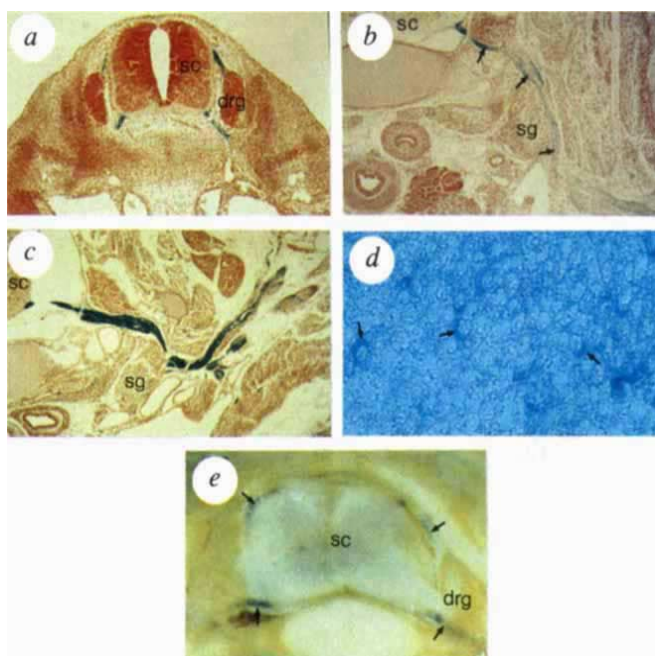


FIG. 1 Expression of *Krox-20* in the developing PNS. a–d, Staining for β -galactosidase activity: a, transverse section through the trunk of an 11.5-d.p.c. *Krox-20*^{-/-} embryo. Note that *Krox-20-lacZ* expression is restricted to the motor and sensory roots of the nerve; b, c, transverse sections through 15.5- and 17.5-d.p.c. *Krox-20*^{+/+} embryos, respectively. At these stages the spinal nerves (arrows) are stained along most of their length. Note the increase in the intensity of staining in the older embryo; d, cross-section of a sciatic nerve from a P10 heterozygous mutant mouse. The cytoplasm of Schwann cells (arrows) but not the axons are labelled. e, *In situ* hybridization of a transverse section through the trunk of a 15.5-d.p.c. embryo with a *Krox-20* cDNA probe⁸ labelled with digoxigenin. Note the labelling along both the motor and sensory roots (arrows). sc, Spinal cord; drg, dorsal root ganglion; sg, sympathetic ganglion. For β -galactosidase revelation, embryos and nerves were processed as described⁴. *In situ* hybridization of 150- μ m vibratome-cut embryo sections were as described⁴. Magnification in a, b, c, e: $\times 12$; in d, $\times 120$.

To investigate a possible role of *Krox-20* in PNS myelination, histological and immunocytochemical analyses were done on sections of sciatic nerves from 15-day postnatal wild-type and *Krox-20*^{-/-} mutant mice. Mutant sciatic nerves presented differences in the morphology of the nuclei and a 1.8 ± 0.16 -fold increase in their density (Fig. 2*a, b*). They also showed a large reduction in major components of compacted myelin, including lipidic components, as indicated by difference in luxol blue staining¹⁰ (data not shown), and major myelin proteins like protein zero (P₀) and myelin basic protein (MBP), as well as the myelin-associated fatty-acid transport protein¹¹ P₂ (Fig. 2*c-h*). In contrast, the level of the early Schwann cell marker¹² S-100 and of an early marker of the myelination process, myelin-associated glycoprotein (MAG), were not strongly affected by the mutation (Fig. 2*i-l*). Heterozygous animals did not show any of the above defects (data not shown). The absence of P₀ and MBP and the persistence of S-100 in homozygous mutant nerves were confirmed by western blotting analysis (Fig. 3*a*). Semi-

quantitative polymerase chain reaction (PCR) analysis of P₀ messenger RNA indicated that its level is reduced by 20–35-fold in *Krox-20*^{-/-} mice (Fig. 3*b*). This suggests that the mutation affects late myelin genes at the transcriptional or mRNA processing levels. Analysis of the level of mRNA for MAG also confirmed the protein data, only showing a slight reduction (2–3-fold) in the mutant (Fig. 3*b*). In the CNS, no defect in myelin formation was observed in *Krox-20*^{-/-} animals, consistent with the absence of *Krox-20* expression in oligodendrocytes (data not shown).

We studied the consequences of *Krox-20* inactivation at the ultrastructural level on sciatic nerves from P15 heterozygous and homozygous mutant mice. Examination of semi-thin Epon sections demonstrated the total absence of myelin in nerves from homozygotes, whereas at the same age heterozygous nerves were fully myelinated (Fig. 4*a, b*). This analysis confirmed the increase in Schwann cell number in mutant nerves, although most Schwann cells seemed to relate normally in a

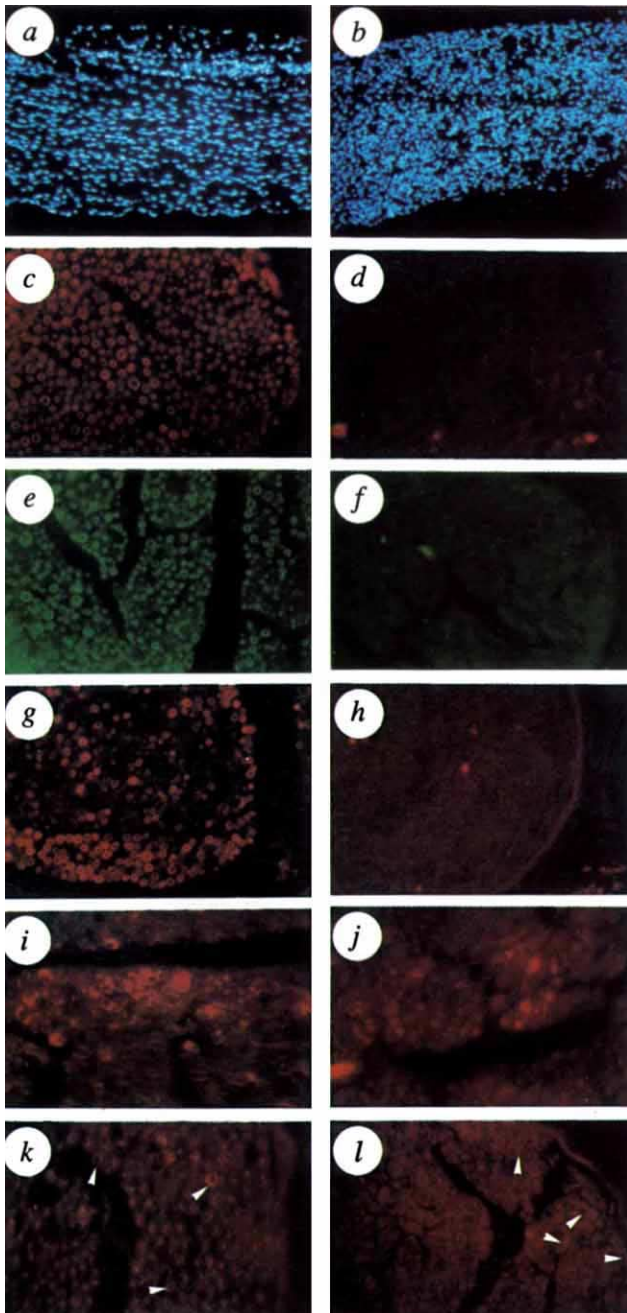


FIG. 2 Histological and immunocytochemical analysis of wild-type and *Krox-20*^{-/-} sciatic nerves. *a, c, e, g, i, k*, Sections through nerves from P15 wild-type mice; *b, d, f, h, j, l*, sections through nerves from P15 homozygous mutant mice. *a, b*, Longitudinal sections stained with bis-benzimide H3358 (Calbiochem) to reveal the nuclei. Note the increase in density and the change in morphology in the mutant: wild-type Schwann cell nuclei present a characteristic flattened shape due to cytoplasmic compression during myelination and are uniformly distributed along the nerve fibres, whereas mutant nuclei are more spherical and are randomly distributed; *c, d*, transverse sections stained with a polyclonal antibody directed against P₀ (gift of B. Trapp); *e, f*, transverse sections stained with a polyclonal antibody directed against MBP (gift of F. Lachapelle); *g, h*, transverse sections stained with a polyclonal antibody directed against P₂ (gift of B. Trapp); *i, j*, transverse sections stained with a polyclonal antibody directed against S-100 (Sigma); *k, l*, transverse sections stained with a polyclonal antibody directed against MAG (gift of B. Trapp). Note the intense crescent-shaped labelling (arrow heads) corresponding to uncompact myelin⁵. In the mutant, the labelling is more diffuse, raising the possibility of a modification of the cellular localization of MAG. Isolated fluorescent cells in *d* and *h* are erythrocytes. Histo- and immunofluorescence were done on paraffin sections as described²³. Magnification in *a, b*, $\times 41$; in *c-l*, $\times 54$.

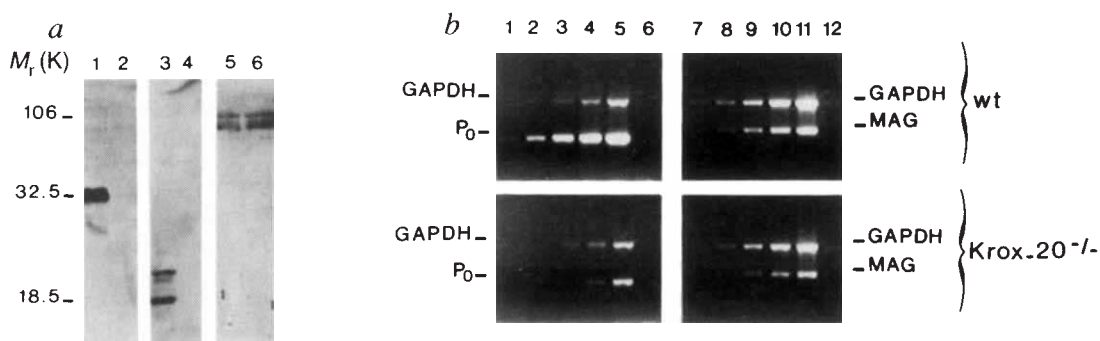


FIG. 3 Analysis of the presence of myelin markers in P10 sciatic nerves. **a**, Western blotting analysis of samples prepared from heterozygous mutant (lanes 1, 3, 5) or *Krox-20*^{-/-} (lanes 2, 4, 6) P10 littermates. The blots were treated with polyclonal antibodies directed against P₀ (lanes 1, 2), MBP (lanes 3, 4) or S-100 (lanes 5, 6). Positions of *M_r* markers are indicated. **b**, Semi-quantitative reverse-PCR detection of P₀ (lanes 1–5) and MAG (lane 7–11) mRNAs prepared from wild-type (wt) and *Krox-20*^{-/-} mice. Amplification of the glyceraldehyde 3-phosphate dehydrogenase (GAPDH) mRNA was done as an internal control. Lanes 6 and 12 represent direct PCR amplification without reverse transcription. The number of amplification cycles was 22 (lane 1), 24 (lanes 2 and 7), 26 (lanes 3 and 8), 28 (lanes 4 and 9), 30 (lanes 5, 6 and 10) or 32 (lanes 11 and 12).

METHODS. Segments of similar length were cut from sciatic nerves. For western blotting, the samples were homogenized directly in sample buffer and boiled for 3 min and further processing was according to

Amersham using the ECL western blotting reagents and a peroxidase-coupled second antibody. For reverse-PCR analysis, RNA was extracted according to Chomczynski and Sacchi²⁴ and reverse-transcribed with an oligodT₁₈ primer. After a 20-fold dilution, PCR amplification was done with Taq polymerase (Perkin Elmer Cetus) under the conditions recommended by the manufacturer. Each cycle of amplification comprised 1 min at 92 °C, 2 min at 65 °C and 2 min at 72 °C. The following primers were used: P₀ 5' primer 5'-GCGCAGCCAGCAGTACCGAATCAG-3' and 3' primer 5'-CTGATTCGGTACTGCTGGCTGCGCA-3'; MAG, 5' primer 5'-GCCACGGTCATCTATGAGAGTCAGC-3' and 3' primer 5'-CAGAATCTCTGGGCACCTGATAAG-3'; GAPDH, 5' primer 5'-TGAAGGTCGGTGTCAACG-GATTGGC-3' and 3' primer 5'-GGTGCCCCAGAGATTCTGAATTCGG-3'. The respective amplified fragments are 477 bp, 604 bp and 982 bp long. After separation on agarose gel and ethidium bromide staining, image acquisition and densitometric analysis were done using a video camera system and the Densylab program (Microvision).

1:1 relationship with axons (Fig. 4*a, b*), suggesting a higher density of the cells along the length of the axons. Electron microscopy analysis confirmed that virtually all axons were devoid of myelin (Fig. 4*c*). Signs of demyelination were not observed, suggesting that the absence of myelin results from an alteration of the onset of the process rather than from myelin degradation. Each axon was surrounded by a Schwann cell that had made about one and a half turns around it (Fig. 4*c*) and had thus initiated myelination. The extracellular space lying between each Schwann cell/axon structure was found to be equivalent in size to the control. This space was filled

with a similar amount of collagen fibrils and each wrapping Schwann cell was surrounded by a basement membrane, excluding the implication of an extracellular matrix defect in the lack of myelination¹³. Therefore in the absence of the *Krox-20* gene product, the myelination process is blocked before the formation of compacted myelin.

Our ultrastructural observations are consistent with the presence of an almost normal amount of MAG in the homozygous mutant nerve and with the absence of P₀ and MBP. Indeed MAG appears early during myelination, is present in uncompacted loops of myelin and is involved in the organization of

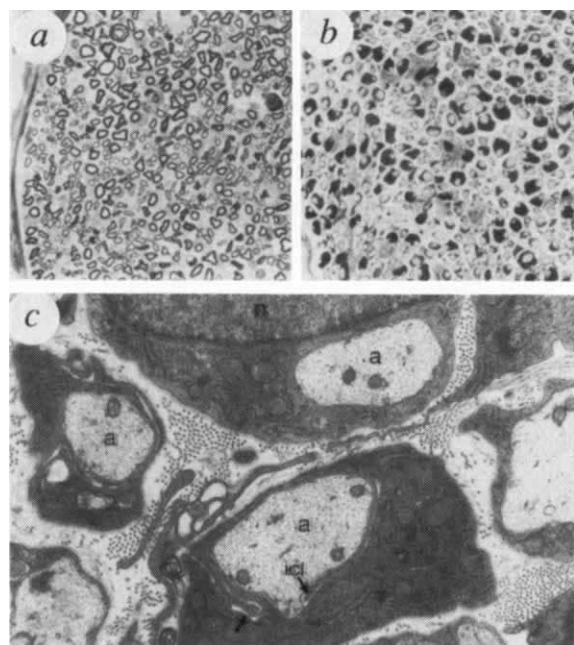


FIG. 4 Cross-sections of Epon-embedded P15 heterozygous and homozygous mutant sciatic nerves. **a, b**, Semi-thin sections from heterozygous and *Krox-20*^{-/-} mice, respectively, photographed under a Leitz light microscope. Note the absence of myelin and the increase in the number of Schwann cell nuclei in the homozygous mutant. **c**, Ultrathin section through a homozygous mutant sciatic nerve photographed under a Jeol electron microscope. Note the absence of compact myelin and that the inner cytoplasmic lip (arrow) of the Schwann cell has progressed past the outer mesaxon (arrow) but has not itself completed a full turn around the axon. All axons were ensheathed. The number of Schwann cells containing more than a single axon was low and similar to the control (data not shown). Homozygous mutant Schwann cell cytoplasm and nucleus are darker than controls. Cytoplasmic organelles such as rough endoplasmic reticulum are prominent and filled with electron-dense material, suggesting high metabolism. **a**, Axon; **n**, nucleus; **icl**, inner cytoplasmic lip. Magnification in **a, b** ×98; in **c**, ×8,250.

the periaxonal space and periaxonal collar^{5,14,16}. In contrast, P₀ and MBP appear later, are present in compacted myelin and are required for myelin compaction^{15,17,20}. The *Krox-20* mutation therefore appears to block Schwann cell differentiation after MAG activation and engulfing of the axon but before spiralization and activation of P₀ and MBP. This phenotype is very similar to the defect observed when differentiating Schwann cells are exposed to antibodies directed against galactocerebroside²¹ (GalC), the major glycolipid of the myelin sheath. It differs, however, from the lesions associated with mutations affecting genes encoding major myelin proteins^{19,20,22}. Our data therefore suggest that *Krox-20* is involved directly or indirectly, in the activation of late myelin genes, although it is not clear whether the absence of late myelin proteins constitutes the cause or the consequence of the differentiation block. Finally, this block is likely to be responsible for the augmentation in the number of Schwann cells by increased proliferation and/or survival. □

Received 20 June; accepted 9 September 1994.

1. Lemke, G. *Glia* **7**, 263–271 (1993).
2. Chavrier, P. et al. *EMBO J.* **7**, 29–35 (1988).
3. Chavrier, P. et al. *EMBO J.* **9**, 1209–1218 (1990).

4. Schneider-Maunoury, S. et al. *Cell* **75**, 1199–1214 (1993).
5. Trapp, B. D. & Quarles, R. H. *J. Cell Biol.* **92**, 877–882 (1982).
6. Greenfield, S., Brostoff, S., Eylar, E. H. & Morell, P. *J. Neurochem.* **20**, 1207–1216 (1973).
7. Kies, M. W., Murphy, J. B. & Alvord, E. C. in *Chemical Pathology of the Nervous System* (ed. Folch-Pi, J.) 197 (Pergamon, London, 1961).
8. Wilkinson, D., Bhatt, S., Chavrier, P., Bravo, R. & Charnay, P. *Nature* **337**, 461–464 (1989).
9. Jessen, K. R. et al. *Neuron* **12**, 509–527 (1994).
10. Gabe, M. *Techniques histologiques* (ed. Masson) 468 (Masson, Paris, 1968).
11. Uyemura, K., Tobari, C. & Hirano, S. *Biochim. biophys. Acta* **214**, 190–197 (1970).
12. Brookes, J. P., Fields, K. L. & Raff, M. C. *Brain Res.* **165**, 105–118 (1979).
13. Bunge, M. B., Bunge, R. P., Kleitman, N. & Dean, A. C. *Dev. Neurosci.* **11**, 348–360 (1989).
14. Trapp, B. D., Quarles, R. H. & Suzuki, K. J. *Cell Biol.* **99**, 594–605 (1984).
15. Trapp, B. D. *J. Cell Biol.* **107**, 675–685 (1988).
16. Li, C. et al. *Nature* **369**, 747–750 (1994).
17. Giese, K. P., Martini, R., Lemke, G., Soriano, P. & Schachner, M. *Cell* **71**, 565–576 (1992).
18. Filbin, M. T., Walsh, F. S., Trapp, B. D., Pizzey, J. A. & Tennekoon, G. I. *Nature* **344**, 871–872 (1990).
19. Kirschner, D. A. & Gansler, A. L. *Nature* **283**, 207–210 (1980).
20. Kimura, M. et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 5661–5665 (1989).
21. Owens, G. C. & Bunge, R. P. *Glia* **3**, 118–124 (1990).
22. Henry, E. W. & Sidman, R. L. *Science* **241**, 344–346 (1988).
23. Levi, G., Crossin, K. & Edelman, G. M. *J. Cell Biol.* **105**, 2359–2372 (1987).
24. Chomczynski, P. & Sacchi, N. *Analyt. Biochem.* **162**, 156–159 (1987).

ACKNOWLEDGEMENTS. P.T. and S.S.-M. contributed equally to this work. We thank B. Trapp for the gift of anti-P₀, anti-MAG and anti-P₂ antibodies, P. Brophy and F. Lachapelle for the gifts of anti-MAG and anti-MBP antibodies, respectively, M. Wassef and C. Sotelo for critical comments on the manuscript, D. Pham-Dinh, P. Cameron-Curry, N. Baumann and B. Zalc for helpful discussions. This work was supported by grants from INSERM, MRT, EEC, ARC, LNFCC, AFM and AISM. The financial support of Telethon (Italy) to the project 'Molecular modulation of cell adhesion and cytoskeletal molecules during neuromuscular development and regeneration' is gratefully acknowledged.

Agouti protein is an antagonist of the melanocyte-stimulating-hormone receptor

Dongsi Lu*, Derril Willard†, Indravadan R. Patel‡, Sue Kadwell†, Laurie Overton†, Tom Kost†, Michael Luther†, Wenbiao Chen*, Richard P. Woychik‡, William O. Wilkison†§ & Roger D. Cone*

* Vollum Institute for Advanced Biomedical Research, Oregon Health Sciences University, Portland, Oregon 97201, USA

† Division of Molecular Sciences, Glaxo Research Institute, Research Triangle Park, North Carolina 27709, USA

‡ Biology Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-8077, USA

THE genetic loci *agouti* and *extension* control the relative amounts of eumelanin (brown-black) and pheomelanin (yellow-red) pigments in mammals¹: *extension* encodes the receptor for melanocyte-stimulating hormone (MSH)² and *agouti* encodes a novel 131-amino-acid protein containing a signal sequence^{3,4}. Agouti, which is produced in the hair follicle⁵, acts on follicular melanocytes⁶ to inhibit α -MSH-induced eumelanin production, resulting in the subterminal band of pheomelanin often visible in mammalian fur. Here we use partially purified agouti protein to demonstrate that agouti is a high-affinity antagonist of the MSH receptor and blocks α -MSH stimulation of adenylyl cyclase, the effector through which α -MSH induces eumelanin synthesis. Agouti was also found to be an antagonist of the melanocortin-4 receptor^{7,8}, a related MSH-binding receptor. Consequently, the obesity caused by ectopic expression of agouti in the lethal yellow (A^y) mouse⁹ may be due to the inhibition of melanocortin receptor(s) outside the hair follicle.

Two models have been proposed for the mechanism by which agouti protein inhibits stimulation of melanogenesis by α -MSH: (1) agouti is a negative regulator of the cyclic AMP signalling pathway acting through a unique agouti receptor¹⁰, or (2) agouti is a competitive antagonist of α -MSH¹¹. We produced recom-

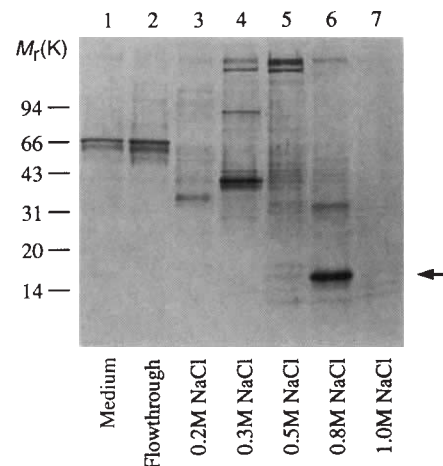


FIG. 1 Production and purification of recombinant agouti polypeptide. A 614-bp *Xba*I/*Pst*I fragment of the full-length mouse agouti cDNA was subcloned into a *Xba*I/*Pst*I-digested baculovirus expression vector pAcMP3 (PharMingen, San Diego). Virus was produced using standard methods²². 1 μ g of each sample was electrophoresed on a 4–20% Tris-glycine gel (Novex, San Diego) and visualized by ProBlue (Integrated Separation Systems, MA) staining. The agouti protein eluted with the 0.8 M-NaCl wash. The 18.5K agouti species (arrow) was not observed in media infected with wild-type virus, and was demonstrated, after elution from the gel, to contain agouti by N-terminal sequencing. Lanes: 1, 48 h post-infection medium from pAcMP3-M.agouti-infected cells; 2, flow-through from Poros-20 HS column; 3–7, NaCl elutions from Poros-20 HS column. Arrow indicates authentic agouti protein.

METHODS. *T. ni* cells were infected at an MOI of 2 with pAcMP3-M.agouti virus or control virus and media were collected 48 h post-infection. This medium was directly loaded onto a Poros-20 HS cation-exchange column (PerSeptive Biosystems, MA) and bound protein eluted with NaCl concentrations from 0.2–1.0 M. The 0.8 M fraction was dialysed into 50 mM NaCl, 20 mM PIPES, pH 6.5, and then diluted for assay. Agouti concentrations were estimated from gel electrophoresis and amino-acid analysis of purified protein. Media controls consisted of unrelated baculovirus supernatants collected 48 h post-infection and purified by NaCl elution of a Poros-20 HS column as for agouti.

§ To whom correspondence should be addressed.

binant agouti protein using the baculovirus expression system in order to test these hypotheses. The band indicated by the arrow in Fig. 1 was absent from *Trichoplusia* cells not infected with the agouti/baculovirus vector, and its identity as the *agouti* gene product was verified by N-terminal sequencing (data not shown). Agouti pooled from the 0.8 M NaCl elution shown in this preparation was estimated to be 75% pure, resulting in an effective agouti concentration of $\sim 0.14 \text{ mg ml}^{-1}$.

Agouti action was examined initially on the B16F10 murine melanoma cell line¹². Agouti (0.7 nM) shifted the half-maximal effective concentration (EC_{50}) for stimulation of adenylyl cyclase by α -MSH in these cells from $1.7 \pm 0.24 \text{ nM}$ to $13.4 \pm 3.3 \text{ nM}$ (data not shown). To characterize agouti action further, we examined the effects of agouti on the MSH receptor (MSH-R) and other G-protein-coupled receptors stably transfected into the human embryonic kidney 293 cell^{8,13,15}. α -MSH had no effect on the cAMP pathway in untransfected 293 cells, demonstrating the absence of endogenous melanocortin receptors in this cell line (Fig. 2a). Furthermore, addition of agouti (0.7 nM) had no effect on the basal adenylyl cyclase levels. Agouti also had no effect on the ability of thyroid-stimulating hormone (TSH) to bind to its receptor and stimulate adenylyl cyclase in TSH-R-transfected 293 cells (Fig. 2b). As the TSH-R couples to the same G protein as do the melanocortin receptors, G_s , this experiment shows that agouti acts upstream of G_s in the cAMP signalling pathway.

In contrast, the same concentration of agouti protein produced a significant shift in the adenylyl cyclase/functional coupling curve of the murine MSH-R expressed in 293 cells (Fig. 3a). Control supernatants from baculovirus-infected cells at the same total protein concentration had no effect. Agouti protein (0.7 nM) increased the EC_{50} for activation of adenylyl cyclase from $1.5 \pm 1.1 \times 10^{-9}$ to $2.2 \pm 1.2 \times 10^{-8} \text{ M}$, but did not alter the maximally induced activity of the receptor. The apparent K_i value calculated from four independent experiments was $3.2 \pm 2.6 \times 10^{-10} \text{ M}$. A dose-response curve demonstrated increasing inhibition of the murine MSH-R by agouti at concentrations from below 10^{-9} M up to 10^{-6} M (Fig. 3b). The human MSH receptor was also inhibited by agouti but only at much higher protein concentrations (Fig. 3c). It is conceivable that,

unlike its murine counterpart, human agouti is a high-affinity antagonist of the human MSH-R. But as the wild-type agouti pigmentation phenotype is not commonly observed in man, it is possible that agouti no longer antagonizes MSH-R function in the human hair follicle.

The ability of agouti to shift the MSH-R functional coupling curve without affecting maximal receptor activation suggested that agouti was acting as a competitive antagonist. To investigate this further, the ability of the protein to compete with another melanocortin peptide for binding to the mouse MSH-R was examined. Adrenocorticotrophic hormone (ACTH), which contains the α -MSH peptide in its first 13 amino acids, can be radiolabelled at Tyr 23 without loss of potency and binds to the same site on the melanocortin receptors as does α -MSH^{7,15,17}. Competition binding experiments were performed with ¹²⁵I-labelled ACTH₁₋₃₉ and the M-3 subclone of the Cloudman murine melanoma cell line. This cell line was used because of the high density of MSH receptors expressed on the plasma membrane (10,000–50,000), in contrast to the MSH-R-transfected 293 cells, which were estimated to express under 1,000 MSH receptors per cell (data not shown). ACTH bound the MSH-R with a K_d of $6.3 \pm 13.3 \times 10^{-8} \text{ M}$, within the range of previously reported values¹² (Fig. 3d). Using a single-site model, agouti protein blocked 50% of specific ACTH binding to the MSH receptor at a concentration of $1.2 \pm 0.7 \text{ ng ml}^{-1}$ ($K_i = IC_{50}$ (half-maximal inhibitory concentration) = $6.6 \pm 3.8 \times 10^{-10} \text{ M}$) (Fig. 3e). Similar results were obtained when a synthetic radiolabelled α -MSH analogue (Nle₄, D-Phe₇- α -MSH) was used as the radiolabelled ligand (data not shown). As a control, baculovirus supernatant from *T. ni* cells infected with a baculovirus construct containing an unrelated gene insert was purified as described for agouti. This protein had no activity in this assay (Fig. 3e) or on α -MSH stimulation of adenylyl cyclase in MSH-R-transfected 293 cells (data not shown).

So far there are five members of the melanocortin receptor family, MC1-R (MSH-R)^{14,16}, MC2-R (ACTH-R)¹⁴, MC3-R^{15,17}, MC4-R^{7,8} and MC5-R^{18,19}. No functions have yet been described for the latter three members of the family, but the MC3-R and MC4-R are expressed primarily in the central nervous system in brain nuclei involved in neuroendocrine and

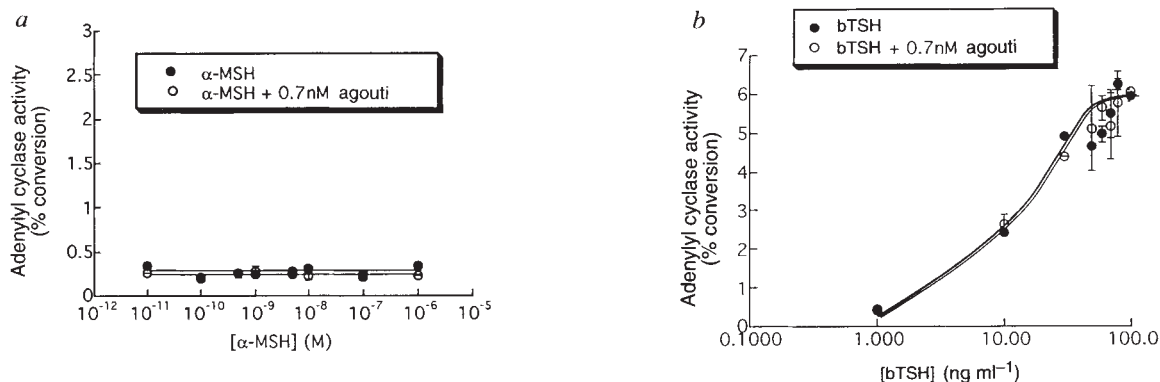


FIG. 2 Agouti does not affect basal or TSH-R-stimulated adenylyl cyclase activity. a, Adenylyl cyclase assay showing no effect of agouti on basal levels of this enzyme in untransfected 293 cells. α -MSH treatment also elicits no response, demonstrating the absence of endogenous melanocortin receptors in this cell line. b, Adenylyl cyclase assay showing no effect of agouti protein on the cAMP signalling pathway following activation by bovine thyroid-stimulating hormone (TSH) of adenylyl cyclase in 293 cells transfected with human TSH-receptor.

METHODS. Agouti protein was prepared as Fig. 1. Two independent preparations of agouti yielded similar results. 293 cells (5×10^5) and 293 cells expressing human TSH-R were preloaded with $5 \mu\text{Ci } ^3\text{H}$ -adenine for 1.5 h at 37°C in a 5% CO_2 incubator. Cells were stimulated with

varying concentrations of α -MSH (a) or bovine TSH (b) in the presence or absence of 0.7 nM agouti for 40 min in cyclase assay incubation medium (Dulbecco's modified Eagle's medium containing 0.1 mg ml⁻¹ bovine serum albumin and 0.1 mM isobutylmethylxanthine). Medium was aspirated and 2.5% perchloric acid containing 0.1 mM cAMP was used to stop the incubation. Adenylyl cyclase activity was calculated by determining the per cent conversion of ^3H -adenine to ^3H -cAMP as described^{23,24}. Correlation of cAMP accumulation measured in this assay with actual adenylyl cyclase activity is made under the assumption that isobutylmethylxanthine effectively blocks cAMP degradation. Data represent means and standard deviations from triplicate data points.

autonomic control. Surprisingly, agouti was also found to be a potent antagonist of α -MSH activation of the MC4-R (Fig. 4a). Once again, the protein appeared to act like a competitive antagonist, not significantly interfering with maximal activation of the receptor. The same concentration of agouti (0.7 nM) did not antagonize activation of the MC3 or MC5 receptors, and the MC5-R was unaffected even at 100 nM agouti concentrations (Fig. 4b, c).

Our results show that agouti is a high-affinity antagonist of the MSH-R, and of at least one of the other melanocortin receptors. Agouti appears to function as a competitive antagonist, inhibiting agonist binding to the MSH-R. This unique bifunctional regulation of the MSH-R by α -MSH and agouti allows for fine spatial and temporal regulation of eumelanin and pheomelanin synthesis. These findings also provide a context for understanding the complex interactions of agouti and extension (MSH-R)

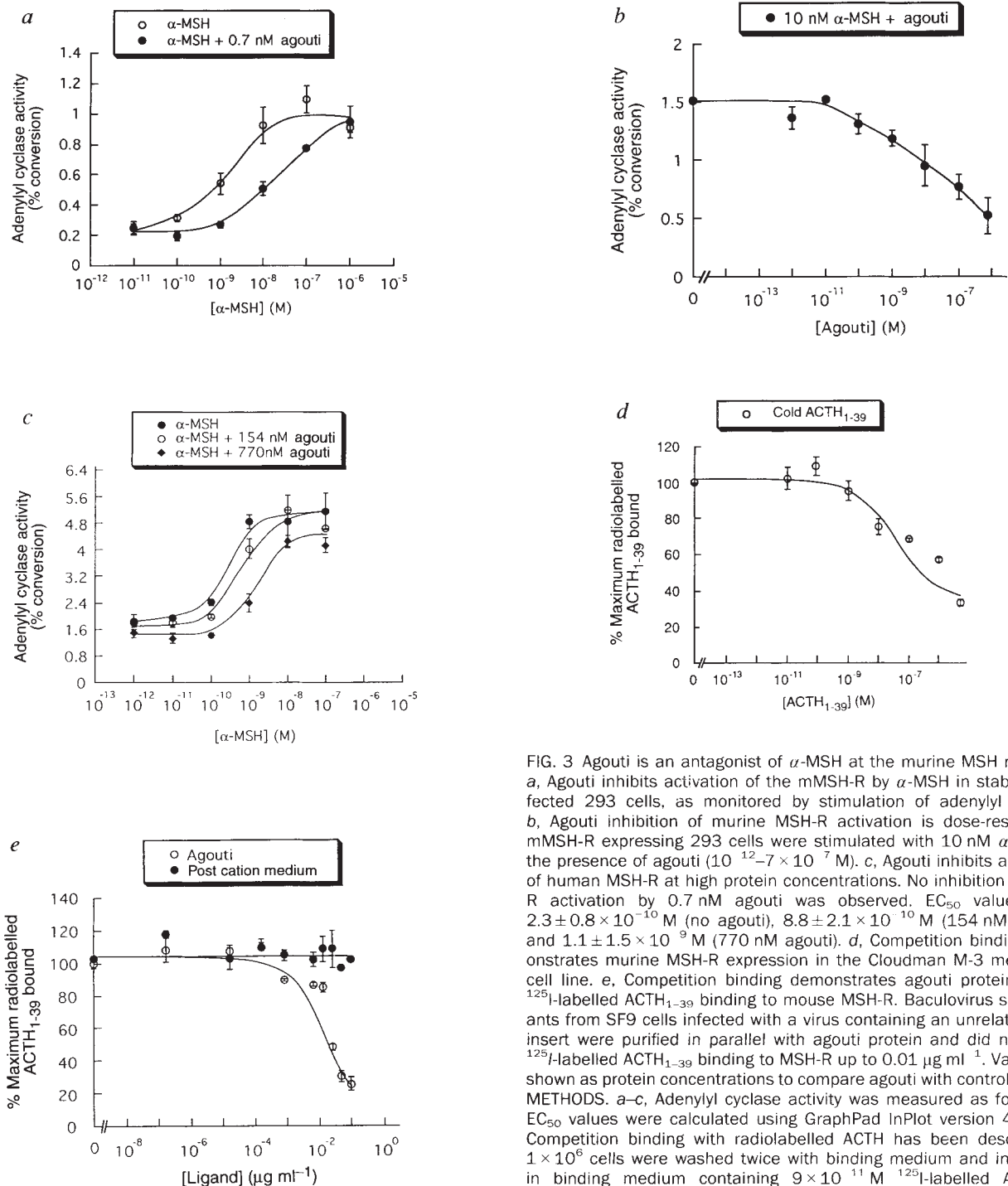


FIG. 3 Agouti is an antagonist of α -MSH at the murine MSH receptor. **a**, Agouti inhibits activation of the mMSH-R by α -MSH in stably transfected 293 cells, as monitored by stimulation of adenylyl cyclase. **b**, Agouti inhibition of murine MSH-R activation is dose-responsive. mMSH-R expressing 293 cells were stimulated with 10 nM α -MSH in the presence of agouti (10^{-12} – 7×10^{-7} M). **c**, Agouti inhibits activation of human MSH-R at high protein concentrations. No inhibition of MSH-R activation by 0.7 nM agouti was observed. EC₅₀ values were $2.3 \pm 0.8 \times 10^{-10}$ M (no agouti), $8.8 \pm 2.1 \times 10^{-10}$ M (154 nM agouti), and $1.1 \pm 1.5 \times 10^{-9}$ M (770 nM agouti). **d**, Competition binding demonstrates murine MSH-R expression in the Cloudman M-3 melanoma cell line. **e**, Competition binding demonstrates agouti protein blocks ¹²⁵I-labelled ACTH₁₋₃₉ binding to mouse MSH-R. Baculovirus supernatants from SF9 cells infected with a virus containing an unrelated gene insert were purified in parallel with agouti protein and did not block ¹²⁵I-labelled ACTH₁₋₃₉ binding to MSH-R up to 0.01 μ g ml⁻¹. Values are shown as protein concentrations to compare agouti with control protein. **METHODS.** **a–c**, Adenylyl cyclase activity was measured as for Fig. 2. EC₅₀ values were calculated using GraphPad InPlot version 4.0. **d, e**, Competition binding with radiolabelled ACTH has been described²⁵: 1×10^6 cells were washed twice with binding medium and incubated in binding medium containing 9×10^{-11} M ¹²⁵I-labelled ACTH₁₋₃₉ (200,000 c.p.m.; Amersham) and different concentrations of agouti, control protein or cold ACTH for 2 h at 22 °C. Cells were then washed, lysed and counted as before²⁵ and data were analysed using the Kaleidagraph software package. Nonspecific binding, determined as the amount of radioactivity bound at 10^{-5} M cold ACTH, was 10% of the total counts bound.

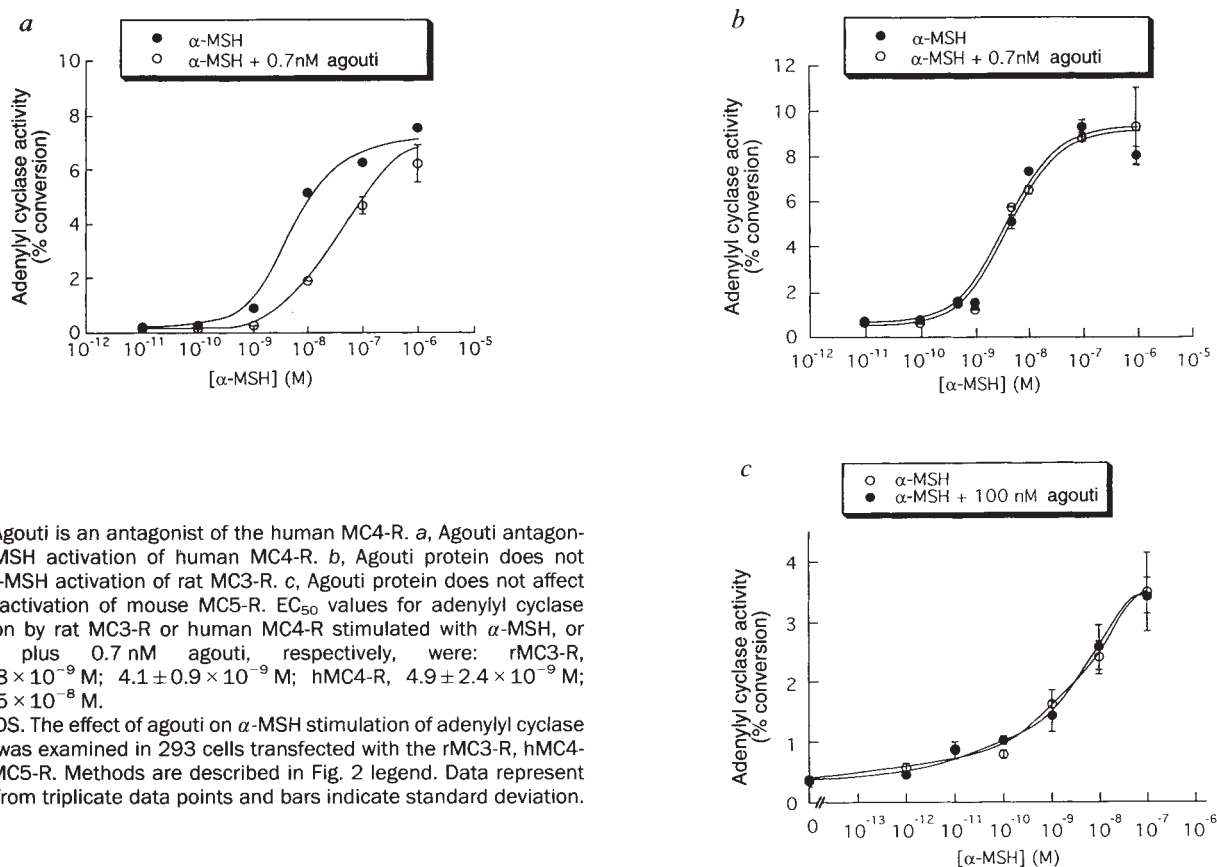


FIG. 4 Agouti is an antagonist of the human MC4-R. *a*, Agouti antagonizes α -MSH activation of human MC4-R. *b*, Agouti protein does not affect α -MSH activation of rat MC3-R. *c*, Agouti protein does not affect α -MSH activation of mouse MC5-R. EC_{50} values for adenylyl cyclase activation by rat MC3-R or human MC4-R stimulated with α -MSH, or α -MSH plus 0.7 nM agouti, respectively, were: rMC3-R, $4.2 \pm 0.8 \times 10^{-9}$ M; $4.1 \pm 0.9 \times 10^{-9}$ M; hMC4-R, $4.9 \pm 2.4 \times 10^{-9}$ M; $3.3 \pm 0.5 \times 10^{-8}$ M.

METHODS. The effect of agouti on α -MSH stimulation of adenylyl cyclase activity was examined in 293 cells transfected with the rMC3-R, hMC4-R or mMC5-R. Methods are described in Fig. 2 legend. Data represent means from triplicate data points and bars indicate standard deviation.

responsible for many mammalian coat colour variants, such as the variable black and tan markings in the German shepherd, resulting from combinations of two *extension* (E^m , E) and three *agouti* (A^s , a^y , a') alleles²⁰.

Because agouti also antagonizes MC4-R function, ectopic overexpression of agouti may lead to obesity in the lethal yellow mouse (A^y) through pathological antagonism of melanocortin receptor(s) expressed outside the hair follicle. Although no agouti pigmentation phenotype has ever been reported in humans, a gene encoding a conserved human agouti protein has been found²¹ and so may have a physiological role. □

Received 13 June; accepted 15 September 1994.

1. Silvers, W. K. *The Coat Colors of Mice: A Model for Mammalian Gene Action and Interaction* (Springer, New York, 1979).
2. Robbins, L. S. et al. *Cell* **72**, 827–834 (1993).
3. Bultman, S. J., Michaud, E. J. & Woychik, R. P. *Cell* **71**, 1195–1204 (1992).
4. Miller, M. W. et al. *Genes Dev.* **7**, 454–467 (1993).
5. Silvers, W. K. & Russel, E. S. *J. exp. Zool.* **130**, 199–220 (1955).
6. Lamoreux, M. L. & Mayer, T. C. *Devl Biol.* **46**, 160–166 (1975).
7. Gantz, I. et al. *J. biol. chem.* **268**, 15174–15179 (1993).
8. Mountjoy, K. G. et al. *Molec. Endocr.* **8**, 1298–1308 (1994).
9. Dickerson, G. E. & Gowen, J. W. *Science* **105**, 496–498 (1947).
10. Conklin, B. R. & Bourne, H. R. *Nature* **364**, 110 (1993).
11. Jackson, I. J. *Nature* **362**, 587–588 (1993).
12. Siegrist, W. et al. *J. Rec. Res.* **8**, 323–343 (1988).
13. Frazier, A. L. et al. *Molec. Endocr.* **4**, 1264–1276 (1990).
14. Mountjoy, K. G. et al. *Science* **257**, 543–546 (1992).
15. Roselli-Rehffuss, L. et al. *Proc. natn. Acad. Sci. U.S.A.* **90**, 8856–8860 (1993).
16. Chhajlani, V. & Wikberg, J. E. S. *FEBS Lett.* **309**, 417–420 (1992).
17. Ganz, I. et al. *J. biol. chem.* **268**, 8246–8250 (1993).
18. Chhajlani, V., Muceniece, R. & Wikberg, J. E. S. *Biochem. biophys. Res. Commun.* **195**, 866–873 (1993).
19. Desarnaud, F. et al. *Biochem. J.* **299**, 367–373 (1994).
20. Little, C. C. *The Inheritance of Coat Colors in Dogs* (Macmillan, New York, 1957).
21. Kwon, H. Y. et al. *Proc. natn. Acad. Sci. U.S.A.* (in press).
22. Summers, M. D. & Smith, G. E. *Texas Agric. exp. St. Bull.* 1555 (1987).
23. Salomon, Y. *Meth. Enz.* **195**, 22–28 (1991).
24. Johnson, R. A. & Salomon, Y. *Meth. Enzym.* **195**, 3–21 (1991).
25. Rainey, W. E., Viard, J. M. & J. M. *J. biol. chem.* **264**, 21474–21477 (1989).

ACKNOWLEDGEMENTS. This work was supported by NIH grants (R.D.C.), and by the Office of Health and Environmental Research, US Department of Energy, under contract with Martin Marietta Energy Systems, Inc. (R.P.W.). We thank J. Weier, P. DeLacey, and K. Lewis for their help and M. Furth and T. Kenakin for discussion.

Inefficient gene transfer by adenovirus vector to cystic fibrosis airway epithelia of mice and humans

Barbara R. Grubb, Raymond J. Pickles, Hong Ye, James R. Yankaskas, Ralph N. Vick, John F. Engelhardt*, James M. Wilson*, Larry G. Johnson & Richard C. Boucher

CF/Pulmonary Research and Treatment Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7020, USA

* Institute of Human Gene Therapy, University of Pennsylvania, Philadelphia, Pennsylvania 19104-4268, USA

THE success of adenoviral vectors for gene therapy of lung disease in cystic fibrosis (CF) depends on efficient transfer of the complementary DNA encoding the correct version of the cystic fibrosis transmembrane regulator (CFTR) to the affected columnar epithelial cells lining the airways of the lung. Pre-clinical studies *in vitro* suggest that low doses of adenovirus vectors carrying this CFTR cDNA can correct defective Cl^- transport in cultured human CF airway epithelia¹. Here we use mice carrying the disrupted CF gene² to test the efficacy of this transfer system *in vivo*. We find that even repeated high doses can only partially (50%) correct the CF defect in Cl^- transport *in vivo* and do not correct the Na^+ transport defect at all. We investigated this discrepancy between the *in vivo* and *in vitro* transfer efficiency using CF mouse

and human samples, and found that it reflects a difference in the susceptibility to adenovirus-5 transduction of the epithelial cell types dosed *in vivo* (columnar) and *in vitro* (basal-cell-like). These studies indicate that more efficient adenoviral gene-transfer vectors and/or refinement of dosing strategies are needed for therapy of CF lung disease.

Nasal mucosae have been used to study epithelial Na^+ and Cl^- transport defects^{3,4} and adenovirus-mediated gene transfer⁵⁻⁷ in CF patients. We adapted the technique used to measure *in vivo* nasal potential difference (p.d.) in humans⁸ to investigate Na^+ and Cl^- transport in the nasal mucosa of mice (Fig. 1A). The basal p.d., an index of Na^+ transport⁹, was raised in CF mice compared with normal mice (Fig. 1A, C). The p.d. response to removal of Cl^- from the luminal perfusate (Cl^- -diffusion p.d.) was used to measure the relative permeability of the epithelium for Cl^- . The normal mouse nasal mucosa exhibited a large hyperpolarization (Fig. 1A, a) in response to this manoeuvre, whereas nasal mucosa depolarized in the CF mouse (Fig. 1A, b). Thus, the *in vivo* p.d. can be used in mice, as in humans, to detect correction (normalization) of CF Na^+ and Cl^- transport defects (Fig. 1A, c, d).

The protocols for testing the efficacy of an Ad5-CB-CFTR vector¹⁰ in CF mice were similar to those used in human clinical trials. As preliminary data indicated that low-dose vector was ineffective, we gave the highest dose of clinical grade Ad5-CB-CFTR^{6,7} that could be achieved (8×10^{11} PFU ml^{-1} delivered in 30 μl to the mouse nasal cavity, corresponding to a multiplicity

of infection (MOI) of 10^4 (ref. 11)) and compared single and multiple dosing regimens (once daily for four days).

A single maximal dose did not restore Cl^- transport in CF mice (Fig. 1B, a, c). Multiple dosing resulted in significant p.d. responses to Cl^- substitution (Fig. 1B, b, c), restoring ~50% of normal Cl^- transport on day 7 (compare Cl^- -diffusion p.d. response (ΔV_i) in Fig. 1B, c with the values for a normal response in Fig. 1A, d). As the efficiency of gene transfer, defined as the percentage of cells corrected in an epithelial surface area, is indicated by restoration of Cl^- transport¹², then efficiency here was low (<10%). Immunohistochemistry (Fig. 2a-f) and *in situ* hybridization (Fig. 2g-j) analysis for CFTR of corrected (day-7) mice¹⁰ detected only patchy (<5% of cells) CFTR expression in the nasal epithelium. Thus, functional and morphological assessment both indicate that *in vivo* Ad5-mediated CFTR gene transfer to mouse nasal mucosa is inefficient.

We tested whether this inefficient gene transfer *in vivo* could correct nasal Na^+ transport in CF mice⁹, and found that Na^+ transport (basal p.d.) did not return to normal values in single or multiple dosing cohorts (Fig. 1B, d), similar to findings from a study of CFTR gene transfer by liposomes into nasal epithelia of CF mice¹³. It is not yet known whether correction of Cl^- without reducing Na^+ hyperabsorption will ameliorate CF lung disease, but our results suggest that inefficient gene transfer may not be adequate to treat CF lung disease¹⁴.

Our results show that *in vivo* Ad5-mediated correction of bioelectric defects in CF mice is much less efficient than that found

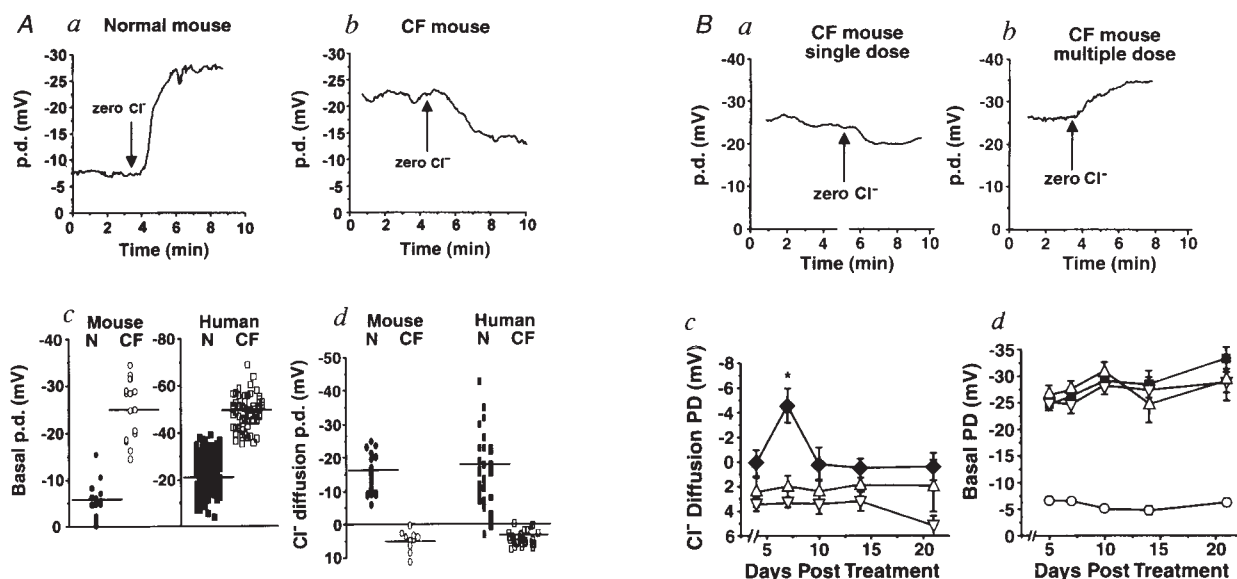


FIG. 1 *In vivo* nasal p.d. in normal and CF mice (A); CF mice nasal bioelectric responses to Ad5-CB-CFTR vector administration (B). A, Tracing of *in vivo* nasal p.d. of normal (a) and CF (b) mouse under basal conditions and in response to luminal Cl^- substitution. c, Summary of basal p.d. measurements in normal (N) and CF mice and humans, and d, summary of p.d. responses to Cl^- substitution for normal and CF mice compared to humans^{3,4} (horizontal bars, mean). B, Recording of nasal p.d. of CF mice dosed once (a) or 4 times (b) with Ad5-CB-CFTR. c, Correction of Cl^- transport (Cl^- -diffusion p.d.) in CF mice dosed with Ad5-CB-CFTR 4 $\times$ (◆), 1 $\times$ (△), and control CF mice (▽). d, Basal p.d.s in normal, control CF, and treated CF mice: Ad5-CB-CFTR 4 $\times$ dose (■), 1 $\times$ dose (△), control CF (▽), normal mice (○). METHODS. CF mice and littermate controls²⁴ (male and female, 2–4 months old) were bred at UNC. The mean body masses of the normals, CF controls, and CFTR-treated CF mice at the beginning of the study were 26.7 ± 1.0 g ($n=15$), 23.6 ± 1.4 g ($n=19$) and 24.5 ± 1.1 g ($n=18$), respectively. After measuring the basal p.d.⁹, the perfusion solution was switched to a low- Cl^- solution (130 mM sodium gluconate replaced

130 mM NaCl), and perfusion continued until a new steady state (usually 5 min) was reached. The Cl^- -diffusion p.d. was calculated as the difference between the basal p.d. and the steady-state p.d. during low Cl^- solution, uncorrected for liquid junction potentials⁴. Data for both nostrils were averaged. The p.d. response to terbutaline (10^{-5} M) of amiloride-pretreated (10^{-4} M) nasal epithelium was also tested in normal and CF mice. Adenoviral vectors containing the *lacZ* (Ad5-CB-*lacZ*)¹⁰ or *CFTR* (Ad5-CB-CFTR) gene^{6,7,10} were administered over a 5-min period to normal and CF mice by instilling vector (in 3% glycerol in PBS, pH 7.4) into the nasal cavity. Virus solution was retained in the cavity for >20 min. The investigator measuring nasal p.d. was blind with respect to treatment. The normal group of mice were Ad5-CB-*lacZ*-dosed mice and sample size ranged from 15–12 mice. Control CF mice (19–8 animals) received no virus or Ad5-CB-*lacZ* (dose 1 $\times$ or 4 $\times$). The number of mice declined during the course of the experiment as a result of anaesthetic mortality. Statistical analyses were performed using analysis of variance and Student Neuman-Keuls test. Asterisk denotes $P < 0.05$.

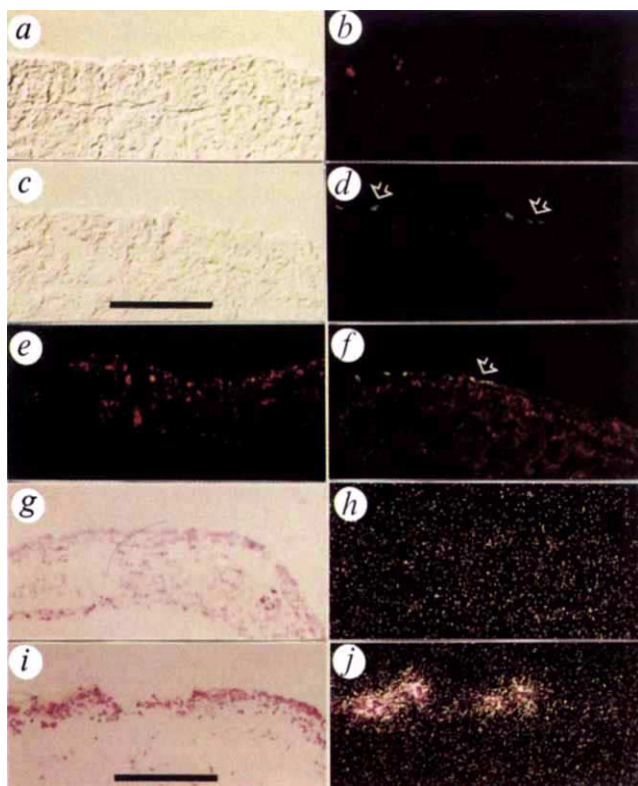


FIG. 2 Detection of recombinant CFTR protein and mRNA in CF mouse nasal epithelium. Indirect immunofluorescence detection of transgene-derived CFTR protein using a polyclonal antibody to human CFTR (amino acids 1,456–1,468) in mouse nasal epithelium infected four times with Ad5-lacZ (a, b and e) and Ad5-CB-CFTR (c, d and f). Photomicrographs show Nomarski images of columnar epithelium (a, c) and fluorescein isothiocyanate fluorescent images (b, d, e and f) *in situ* hybridization detected transgene-derived CFTR mRNA using a human CFTR R-domain antisense riboprobe within epithelium infected four times with Ad5-lacZ (g, h) and Ad5-CB-CFTR (i, j). Brightfield (g, i) and darkfield (h, j) photomicrographs are shown. Immunofluorescent and *in situ* hybridization detection of transgene-derived CFTR was done in serial sections. Epithelium infected with Ad5-lacZ showed no signal with sense and RNase-treated antisense probes (data not shown). a–d, Scale bar, 80 μ m; in e–j, scale bar, 160 μ m.

METHODS. Nasal epithelium infected with Ad5-lacZ and Ad5-CB-CFTR were excised from animals on day 7. Immediately following *in vivo* transepithelial p.d. measurement, the tissues were snap-frozen in OCT-embedding compound. Fresh frozen 6- μ m sections were processed for detection of CFTR protein by immunohistochemistry and CFTR mRNA by *in situ* hybridization as described²⁵.

in a study describing highly efficient (MOI, 1–25) Ad2-mediated correction of nasal bioelectric abnormalities in CF patients⁵. To investigate this apparent discrepancy, we first tested whether it was appropriate to measure gene transfer into mice using human Ad5 vectors by determining relative efficiency of LacZ expression and whether human CFTR could correct defective epithelia in CF mice. Ad5 expressed LacZ in mouse and human cultures efficiently and equally over a wide dose range (Fig. 3A). Similarly, Ad5-mediated (MOI, 10^1 – 10^2) human CFTR expression restored forskolin/cAMP-regulated Cl^- transport comparably in cultured mouse and human CF epithelia (Fig. 3B). Figure 3C and D compares the morphology of mouse and human cultured cells and the efficiency of CFTR transduction. Cells from both species were cuboid in shape (Fig. 3C, a and D, a) and expressed keratin-14 (K14; Fig. 3C, b and D, b), features characteristic of basal-cell-like airway epithelial cells¹⁵. Human CFTR (and LacZ; result not shown) was expressed in virtually all cells (>90%) in both mouse (Fig. 3C, c) and human (Fig. 3D, c) monolayers.

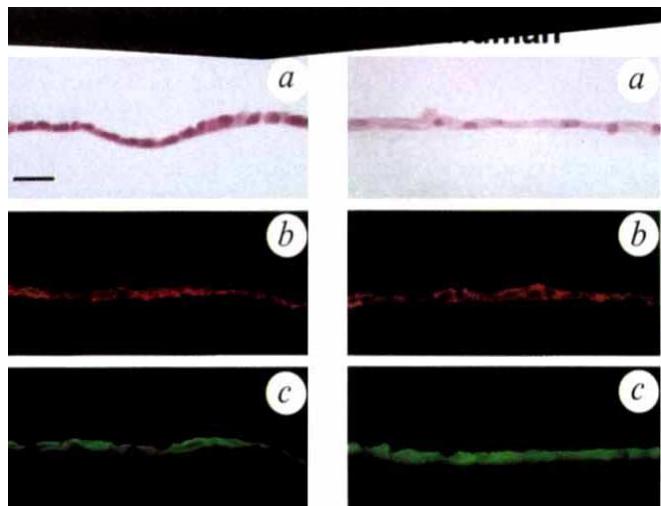
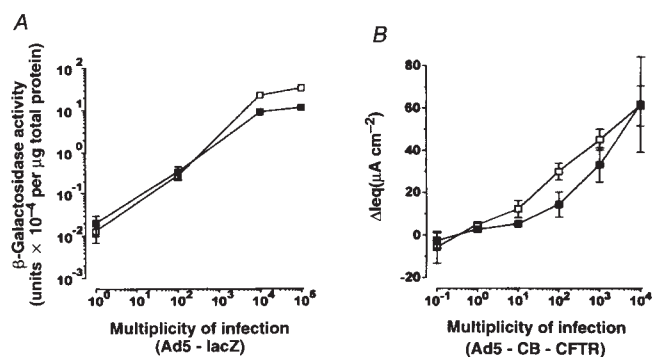


FIG. 3 Comparison of the transgene expression in murine and human nasal epithelial cell primary cultures. Ad5-lacZ-mediated (A) and Ad5-CB-CFTR-mediated gene transfer (B) are compared in cultures of mouse (□) and human (■) nasal epithelia. β -Galactosidase activity was used to quantify LacZ (A) and forskolin-stimulated Cl^- secretion (ΔI_{eq}) to quantitate CFTR. C, a and D, a, Photomicrographs of murine and human cell monolayers, respectively, at the time of Ad5 vector infection (H&E stained). C, b and D, b, Immunofluorescent detection of K14 (phycoerythrin fluorescence) in cultured airway epithelia. No fluorescent signal was detected when either the primary anti-K14 antibody was omitted or an irrelevant primary antibody was used (data not shown). Representative sections of Ad5-CFTR-treated murine (C, c) and human (D, c) cell monolayers after exposure to Ad5-CB-CFTR (MOI, 10^4), followed at 48 h by indirect immunofluorescent detection of CFTR protein using fluorescein isothiocyanate. Control is described in Fig. 2 legend. Scale bar in C, D, 40 μ m.

METHODS. Nasal epithelial cells were isolated from excised nasal turbinates and cultured on permeable collagen supports in serum-free conditions^{24,26}. CF cells were used for Ad5-CFTR experiments and normal cells for those involving Ad5-lacZ. Cell cultures were exposed to Ad5-lacZ for 30 min at 37 °C, washed and incubated for a further 48 h before analysis. For functional analysis of CFTR, cells at confluence were infected (18 h) with Ad5-CFTR. Bioelectrical studies were done after 72 h as described¹². Each point in A and B represents the mean $\pm$ s.e.m. ($n \geq 3$ per point). For indirect immunofluorescent detection, frozen sections of monolayers were used²⁵; CFTR protein was detected as described for Fig. 2, K14 was detected using a polyclonal anti-mouse K14 antibody and a monoclonal anti-human K14 antibody (Vector Labs; modified from ref. 27).

We next compared *in vivo* gene transfer in mouse and human airways using tracheae of mice, dosed intraluminally with Ad5-lacZ, and human tracheal specimens (a model for the well differentiated columnar epithelium characteristic of human patients), dosed with Ad5-lacZ immediately after surgical removal from transplant patients. Because transduction efficiency as assessed by LacZ histochemistry was sparse (<1%), expression was compared 24 h after dosing using β -galactosidase assay

(Fig. 4A, B). Expression of LacZ in murine tracheal epithelia *in vivo* was similar to that in freshly excised human tissue (compare Fig. 4A, B). Efficiency in both species (Fig. 4A, B), however, differed markedly *in vivo* and *in vitro*.

The basis for this difference is unlikely to be immunological, as this would limit duration rather than the extent of expression¹⁶, but it might reflect a difference in susceptibility to Ad5 transduction of the different cell types exposed to vector *in vitro* (basal-cell-like phenotype; Fig. 3C, D) and *in vivo* (columnar cells, both ciliated and secretory types; Fig. 2). We therefore removed well differentiated columnar epithelial cells by abrasion, exposing intermittent stretches of basal cells (as well as the adjacent uninjured columnar epithelium) to luminal Ad5-*lacZ*. Both murine (Fig. 4C, a) and human airways (Fig. 4D, a) showed evidence of preferential Ad5-*lacZ* gene transfer in the regions of injury. LacZ was expressed in cuboidal cells in these areas but not in adjacent columnar epithelium (Fig. 4C, b and D, b). Immunohistochemistry showed that cells expressing the transgene (Fig. 4C, c and D, c) also expressed the basal-cell marker, K14 (ref. 15). Thus, the difference between gene-transfer efficiency *in vivo* and *in vitro* reflects the dosing *in vivo* of columnar cells, which are poorly transducible by adenovirus vectors, whereas airway epithelia dosed *in vitro* have acquired the more transducible basal-cell-like phenotype (Fig. 3).

We conclude that the discrepancy between our results and those of the human CF trial⁵ does not reflect a species difference response to adenovirus-mediated gene transfer. No difference arises from the use of Ad2 rather than Ad5 as vector, because in cultured human CF monolayers the correction of forskolin-stimulated Cl^- currents is comparable for both vectors (data not shown). The discrepancy could be a result of a difference in the site of vector administration: in the mouse trial, vector was deposited on columnar respiratory epithelium (Fig. 2)¹⁷; in the human trial it was applied on the medial surface of the inferior turbinate, an area that may be more susceptible to gene transfer as it has cuboidal morphology, which reflects continual epithelial damage and repair at this site⁸. Alternatively, the *in vivo* protocol employed in the human trial⁵ (the change in potential difference, $\Delta p.d.$, induced by terbutaline in nasal mucosa pre-treated with amiloride), which in our experience does not discriminate between normal and CF mice ($\Delta p.d. = 0.8 \pm 0.4$ mV, $n=8$ for normal mice; -0.3 ± 2.0 mV, $n=8$ for CF mice) or between normal and CF humans ($\Delta p.d. = -0.4 \pm 1.5$ mV, $n=6$ for normal subjects; -0.3 ± 0.7 mV, $n=8$ for CF subjects (M. R. Knowles, personal communication)), may not have been adequate to detect a corrective effect.

Successful gene transfer by an adenovirus vector to differentiated columnar-type cells has been reported *in vivo*¹⁸, which may in part reflect transduction of basal cells that have differentiated into columnar cells¹⁵ over the 24–72 h between vector administration and histochemical assay of LacZ expression. Alternatively, the differences may be a result of regional variation in respiratory tract epithelia. Most *in vivo* studies reporting efficient adenovirus-mediated LacZ expression have focused on rodent lower airways, whereas our mouse studies were on the upper airways, which have a distribution of epithelial cell types different from that of mouse lower airways and more like that of humans¹⁷. Our data from human specimens (Fig. 4D) are congruent with studies of non-human primates^{19,20}, in which adenovirus-mediated *lacZ* transfer in the airways is not efficient.

Our results demonstrate that Ad5-mediated gene transfer into murine nasal and tracheal mucosa *in vivo* and into freshly excised human specimens is poor and highlight the need to define the status of the epithelium targeted for gene transfer in pre-clinical and clinical studies. The requirement for high concentrations of Ad5 vector for successful transfer indicates that safety concerns^{21,22} could be relevant to human gene therapy trials. The transfer efficiency of adenovirus vectors needs to be improved and delivery strategies refined, for example by increasing access to basal cells, which will enhance gene transfer in the airway,

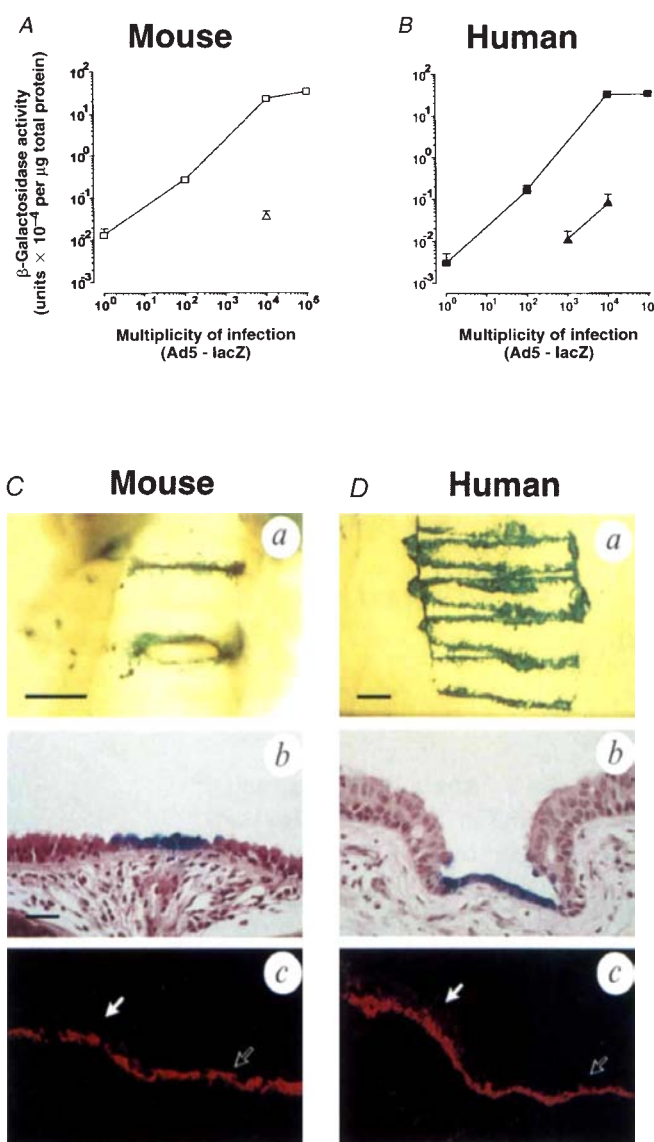


FIG. 4 *In vivo* versus *in vitro* Ad-*lacZ* gene transfer efficiency (A, B) and transduction efficiency in injured versus uninjured regions of airway epithelium (C, D). A, Comparison of LacZ expression, quantified as β -galactosidase activity, mediated by Ad5-*lacZ* in cultured ($\square$) and *in vivo* (Δ) mouse airway epithelia. B, Data for human CF tracheae, cultured ($\blacksquare$) versus acutely dosed freshly excised tissue ($\blacktriangle$). *En face* view through a dissecting microscope of LacZ expression in abraded (linear regions) murine trachea (C, a) and abraded freshly excised human CF trachea (linear regions; D, a). Histological sections perpendicular to linear regions of injury and stained with haematoxylin and eosin are shown in C, b, D, b. Indirect immunofluorescent detection of K14, in parallel tissues, are shown in C, c and D, c. Damaged and undamaged tissue from both species not exposed to Ad5-*lacZ* do not express LacZ (data not shown). Control for K14 staining was as for Fig. 3. Scale bars: in C, a and D, a, 200 μm ; in C, b, and D, b, 30 μm .

METHODS. Tracheae of anaesthetized mice were exposed and proximal and distal tracheostomies performed. Viral solution (10 μl) was instilled into the trachea where it remained for 30 min; then it was evacuated from the trachea, and the animal allowed to recover until it was killed 24 h later. Human CF tracheal specimens ($n=3$ patients) were cut into 1-cm² samples, exposed to Ad5-*lacZ* for 30 min; tracheae were then rinsed, and incubated²⁸ for 24 h before processing. Cells from mouse and human airways were collected and cultured as described in Fig. 3 legend. At MOI of 10^4 , vector concentrations were: A, mouse, cultured 1.5×10^9 PFU ml^{-1} and *in vivo* 1×10^{12} PFU ml^{-1} ; B, human, cultured 1.5×10^9 PFU ml^{-1} and *in vivo*, 4×10^{11} PFU ml^{-1} . Enzymatic assay for LacZ was performed on cultured and freshly isolated airway epithelial cells²⁶ as described for Fig. 3 ($n \geq 3$ per point). Injury was by mechanical abrasion of linear regions of epithelium using forceps; areas avoiding forceps contact were not damaged.

especially in young CF patients in whom airway injury is minimal. The CF mouse seems to be an accurate and useful model^{13,23} for testing such strategies. □

Received 30 June; accepted 9 September 1994.

1. Zabner, J., Couture, L. A., Smith, A. E. & Welsh, M. J. *Hum. Gene Ther.* **5**, 585–593 (1994).
2. Snouwaert, J. et al. *Science* **257**, 1083–1088 (1992).
3. Knowles, M. R., Gatz, J. & Boucher, R. *New Engl. J. Med.* **305**, 1489–1495 (1981).
4. Knowles, M. R., Gatz, J. & Boucher, R. *J. clin. Invest.* **71**, 1410–1417 (1983).
5. Zabner, J. et al. *Cell* **75**, 207–216 (1993).
6. Boucher, R. C. et al. *Hum. Gene Ther.* **5**, 615–639 (1994).
7. Wilson, J. M., Engelhardt, J. F., Grossman, M., Simon, R. H. & Yang, Y. *Hum. Gene Ther.* **5**, 501–519 (1994).
8. Knowles, M. R., Carson, J. L., Collier, A. M., Gatz, J. T. & Boucher, R. C. *Am. Rev. Respir. Dis.* **124**, 484–490 (1981).
9. Grubb, B., Vick, R. & Boucher, R. C. *Am. J. Physiol.* **266**, C1478–C1483 (1994).
10. Engelhardt, J. F. et al. *Nature Genet.* **4**, 27–34 (1993).
11. Gross, E. A., Swenberg, J. A., Fields, S. & Popp, J. A. *J. Anat.* **135**, 83–88 (1982).
12. Johnson, L. G. et al. *Nature Genet.* **2**, 21–25 (1992).
13. Alton, E. W. F. W. et al. *Nature Genet.* **5**, 135–142 (1993).
14. Cutting, G. R. *Nature Genet.* **2**, 4–5 (1992).

15. Shimizu, T., Nettekheim, P., Ramaekers, F. C. S. & Randell, S. H. *Am. J. Respir. Cell molec. Biol.* **7**, 30–41 (1992).
16. Yang, Y. et al. *Nature Genet.* **7**, 362–369 (1994).
17. Harkema, J. R., Mariassy, A., St George, J., Hyde, D. M. & Plopper, C. G. in *The Airway Epithelium. Physiology, Pathophysiology and Pharmacology* (eds Farmer, S. G. & Hay, D. W. P.) 3–39 (Dekker, New York, 1991).
18. Mastrangeli, A. et al. *J. clin. Invest.* **91**, 225–234 (1993).
19. Engelhardt, J. F. et al. *Hum. Gene Ther.* **4**, 759–769 (1993).
20. Bout, A. et al. *Hum. Gene Ther.* **5**, 3–10 (1994).
21. Simon, R. H. et al. *Hum. Gene Ther.* **4**, 771–780 (1993).
22. Zabner, J. et al. *Nature Genet.* **6**, 75–83 (1994).
23. Hyde, S. C. et al. *Nature* **362**, 250–255 (1993).
24. Clarke, L. L. et al. *Science* **257**, 1125–1128 (1992).
25. Engelhardt, J. F. et al. *Nature Genet.* **2**, 240–247 (1992).
26. Yankaskas, J. R., Cotton, C. U., Knowles, M. R., Gatz, J. T. & Boucher, R. C. *Am. Rev. Respir. Dis.* **132**, 1281–1287 (1985).
27. Doatney, A. M. et al. *Cell Growth Differentiation* **4**, 1071–1082 (1993).
28. Boman, T. F., Kleiner, J. L., Carlson, D. M., Maloney, W. H. & Matthews, L. W. *Am. Rev. Respir. Dis.* **110**, 428–435 (1974).

ACKNOWLEDGEMENTS. B.R.G. and R.J.P. contributed equally to this project. We thank B. Koller for mice; D. Roop and J. Cohn for anti-mouse K14 and anti-CFTR antisera, respectively; V. Harris, L. Dudus, K. Brigman and S. Boyle for technical assistance; and S. Randell for critical review of the manuscript. Supported by grants from the NIH and CFF.

The *Trithorax-like* gene encodes the *Drosophila* GAGA factor

Gabriella Farkas*, Janos Gausz†, Mireille Galloni*†, Günter Reuter§, Henrik Gyurkovics†|| & François Karch*||

* Department of Zoology and Animal Biology, University of Geneva, 154 route de Malagnou, 1224 Geneva, Switzerland

† Institute of Genetics, Biological Research Center, Hungarian Academy of Sciences, PO Box 521, 6701 Szeged, Hungary

§ Institute of Genetics, Martin-Luther-University, Domplatz 1, D-06108 Halle/S, Germany

LITTLE is known about the way higher-order chromatin structure influences gene expression and chromosome topology in general. Genetic analysis in *Drosophila* has led to the discovery of two classes of genes, the regulators of homeotic genes and the modifiers of position-effect variegation, which seem to be good candidates for encoding some of the factors regulating chromatin functions^{1,2}. The *Trithorax-like* gene we describe here is required for the normal expression of the homeotic genes and is a modifier of position-effect variegation. We found that *Trithorax-like* encodes the GAGA factor which is involved in the formation of an accessible chromatin structure at promoter sequences³. Our genetic analysis suggests that the chromatin modelling function of the GAGA factor is not restricted to promoter regions.

Flies homozygous for the *Trithorax-like*^{13C} (*Trl*^{13C}) mutation have reduced viability, a rough-eye phenotype and are female-sterile. On the basis of its interaction with *Ubx*, the *Trithorax-like* locus belongs to the *trithorax* group of genes, which are thought to be required for the expression of homeotic genes, such as *Ubx*^{4–6}. In flies doubly heterozygous for *Trl* and *Ubx* (*Trl*^{13C}/+ *Ubx*) halteres are larger than in *Ubx*/+ flies, indicating that the expression of the wild-type *Ubx* gene is reduced. Moreover, at a low frequency (3.9%), patches on the haltere and postnotum develop like wing and notal tissues, respectively, suggesting a complete inactivation of the *Ubx* gene (Fig. 1a–d). Like *trithorax*⁵, *Trl* is also a positive regulator of the *Abd-B* gene of the bithorax complex. Surviving adult males homozygous for *Trl*^{13C} have a few bristles on the sixth sternite, indicating a partial

loss of the *Abd-B*⁺ function in the sixth abdominal segment⁷ (Fig. 1e,f).

Trl^{13C} is also a dominant enhancer of position-effect variegation (PEV). PEV is observed when a euchromatic gene (for example, *white*) is brought into the vicinity of heterochromatin by a chromosomal rearrangement. In this situation heterochromatin may inactivate the *white* gene in a clonally heritable manner, giving rise to a mottled eye colour (Fig. 2, right side; for review see ref. 2). The eye on the left side of Fig. 2 is almost completely white, indicating that the *Trl*^{13C} mutation in heterozygous conditions promotes heterochromatinization. Extensive genetic analysis has identified numerous dominant modifiers of PEV. Whereas suppressor loci are thought to be involved in chromatin condensation, genes that mutate to enhancers of PEV, such as *Trl*, might encode chromatin components involved in chromatin decondensation².

The chromosome carrying the *Trl*^{13C} mutation contains a single insertion of a P transposon at the cytological position 70F1-2. Molecular analysis revealed that the insertion occurred in a transcription unit (Fig. 3). To demonstrate that both the *Trithorax-like* and the enhancer-of-PEV phenotypes are due to the same insertion we mobilized the element. Precise excisions resulted in the complete reversion of both phenotypes and restored full viability and fertility as well. We also recovered two deletion derivatives, *Trl*^{R67} and *Trl*^{R85}, which die as homozygotes during the third larval instar with no apparent cuticular phenotype. Because the *Trl*^{R85} deletion removes extensive regions of the transcription unit, we presume that it represents a null allele of the *Trl* locus.

We also localized the insertion site of another P transposon in the 5' upstream region of the *Trl* locus. This mutant (line 62 (ref. 8), referred to here as *Trl*⁶²) was recovered as a moderate dominant enhancer of PEV and appears to be allelic to *Trl*^{13C}. *Trl*^{13C}/*Trl*⁶² flies eclose late, are poorly viable and exhibit the same phenotypes as *Trl*^{13C} homozygotes. *Trl*⁶² homozygotes are late third-instar larval lethal. In mutant larvae, the primitive gonads are undetectable and imaginal discs, especially wing and leg discs, are smaller than normal.

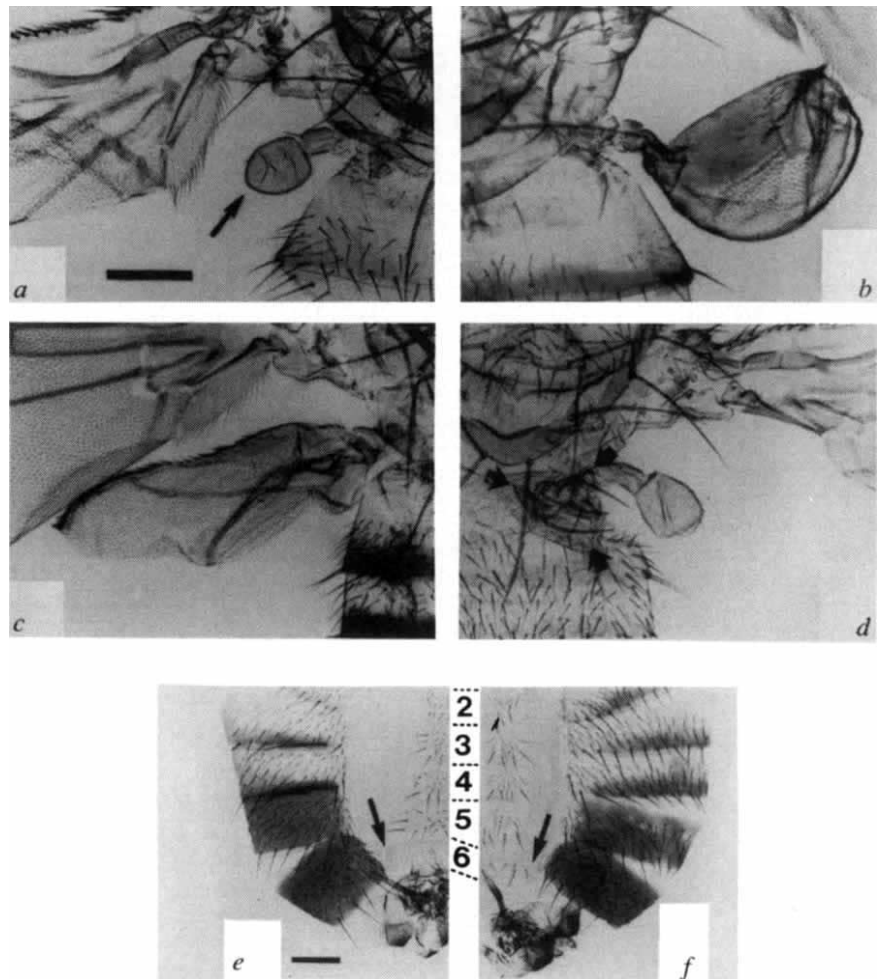
We sequenced the largest of our complementary DNA clones and found that the predicted amino-acid sequence was >99% identical to that of the GAGA transcription factor⁹. We therefore conclude that the *Trl* locus is coding for the GAGA factor.

The GAGA factor has been identified as a protein that binds in a site-dependent manner to the promoter of *Ubx* and stimulates its transcription *in vitro*¹⁰. This agrees with our finding that the *Trl* function is required for the normal expression of the *Ubx* gene. GAGA binding sites are found in the promoter regions of several other *Drosophila* genes^{11–19}. The range of phenotypes associated with different *Trl* alleles suggests that the list

† Present address: Division of Basic Sciences, Fred Hutchinson Cancer Research Center, 1124 Columbia Street, Seattle, Washington 98104, USA.

|| To whom correspondence should be addressed.

FIG. 1 Homeotic transformation in Trl^{13C} mutants. *a-d*, Same magnification. Scale bar, 0.2 mm. *a*, Halteres of flies heterozygous for Ubx mutation (indicated by the arrow) are slightly enlarged as compared to wild type ($Ubx^{130}/+$). *b, c*, Flies heterozygous for both Trl^{13C} and Ubx^{130} mutations have larger halteres than flies carrying the Ubx^{130} mutation alone. Occasionally, patches of wing tissue can be seen in the halteres indicating a nearly complete loss of Ubx expression in these cells (*b*). In some cases, most of the haltere is transformed into wing (*c*). *d*, In the fly shown here, a patch of dorsal tissue in T3 (indicated by the three arrows) adopted the fate of notal tissue, as revealed by the presence of large bristles normally found in the dorsal part of T2. Transformations shown in *b, c* and *d* are not observed in $Ubx^{130}/+$. *e, f*, Whole mount of abdominal cuticle of wild-type (*e*) and homozygous Trl^{13C} males (*f*) were mounted as described in ref. 25. The dorsal midline is at the periphery. Each abdominal segment has a patch of cuticle on the ventral side called sternite. In wild-type males (*e*), the 2nd to 5th sternites are covered with bristles whereas the 6th sternite (arrow) is devoid of hairs. In the Trl^{13C} male (*f*), the 6th sternite has a few bristles (arrow), indicating a homeotic transformation on the 6th abdominal segment into the 5th. This transformation is even stronger in Trl^{13C} hemizygotes, which die just before hatching ($Trl^{13C}/Df(3L)fzM21$; data not shown). Abdominal segments are numbered between the two panels. Scale bar (*e*), 0.2 mm.



of genes regulated by the GAGA factor is probably far from exhaustive.

Initially, the GAGA factor was regarded as a more or less standard positive transcription factor^{10,12,14}. More recently, *in vivo* and *in vitro* studies showed that the GAGA factor enhances transcription by acting as an 'antirepressor'²⁰ and generating nucleosome-free regions of DNA (nuclease-hypersensitive sites) at promoters^{3,19}.

Our genetic analysis suggests that the activity of the GAGA factor is not limited to promoter regions. First, although *Abd-B* is required for the normal development of posterior segments A5 to A8, in Trl^{13C} homozygotes the sixth abdominal segment is transformed most conspicuously (Fig. 1f). This result indicates that the *cis*-regulatory region specific for this segment (*iab-6*), rather than the *Abd-B* gene itself, is affected by the *Trl* mutation. This raises the possibility that the GAGA factor functions not only in opening up promoters, but is also involved in initiating and/or maintaining an active chromatin configuration of distant regulatory elements such as *iab-6*. Second, we find that mutations in the GAGA gene are dominant enhancers of PEV. The effect on PEV and the presence of GAGA protein at a large number of euchromatic sites in polytene chromosomes leads to the idea that the organization of nucleosome-free regions by the GAGA factor may determine the binding of other factors involved in maintaining the boundaries of higher-order chromatin structure. Although *Trl* is the first enhancer of PEV whose gene product is actually shown to affect chromatin structure, the dominant enhancement of PEV may be indirect and mediated



FIG. 2 The Trl^{13C} mutation is a dominant enhancer of position-effect variegation. On the right, the eye phenotype of an $In(1)w^{m4n}$ male is shown. Red spots are clones of cells expressing the *white* gene and white spots are clones of cells in which the *white* gene has been inactivated by heterochromatinization (see text). The eye shown on the left is from an $In(1)w^{m4n}$ male which is heterozygous for the Trl^{13C} mutation. The eye colour is almost white, indicating that the *white* gene is more frequently inactivated by heterochromatinization, a characteristic effect of enhancer mutations². The enhancer effect was confirmed by comparing the eye pigment in a suppressor background, as described in ref. 24 ($Su/+$, 100%; Su/Trl^{13C} , 20%).

through a negative effect on the expression of other enhancer loci.

The GAGA factor contains a putative zinc-finger DNA-binding motif and shares a 120-amino-acid N-terminal region of significant homology, the *tramtrack* motif, with other *Drosophila* trans-acting factors^{16,21,22} and with *kelch*, a component of intercellular bridges in *Drosophila* egg chambers²³. The hydrophobic character of this domain suggests that it may be involved in interactions with other proteins rather than in DNA binding. *E(var)3-93D*, the only other enhancer of PEV whose product is known, also shares homology with the *tramtrack* domain²⁴. Like GAGA mutations, *E(var)3-93D* displays homeotic effects²⁴. In addition, both *Trl^{13C}* and *E(var)3-93D* show interactions with other *trithorax*-group mutations (G.R., G. Hoffmann, J.G. and H.G., manuscript in preparation), raising the possibility that proteins coded by these genes recruit other factors, perhaps through the *tramtrack* domain, to form complexes involved in chromatin modelling. Pointing to the same direction is the observation that the combination of limiting amounts of functional GAGA factor and *Ubx* products in *Trl^{13C}*+/+ *Ubx* trans-heterozygotes occasionally leads to the complete inactivation of *Ubx* which is then clonally inherited. This suggests that GAGA is interacting with factors responsible for transmitting active conformation to the *Ubx* promoter during successive cellular divisions. Whether this interaction is direct remains to be

determined. With the mutations now available it will be possible to identify cofactors by genetic screens. □

Received 3 June; accepted 12 September 1994.

1. Paro, R. *Trends Genet.* **6**, 416–421 (1990).
2. Reuter, G. & Spierer, P. *BioEssays* **14**, 605–612 (1992).
3. Tsukiyama, T., Becker, P. & Wu, C. *Nature* **367**, 525–532 (1994).
4. Lewis, E. B. *Nature* **276**, 565–570 (1978).
5. Ingham, P. W. *Wilhelm Roux Arch. dev. Biol.* **190**, 365–369 (1981).
6. Shearn, A. *Genetics* **121**, 517–525 (1989).
7. Celniker, S. E., Sharma, S., Keelan, D. J. & Lewis, E. B. *EMBO J.* **9**, 4277–4286 (1990).
8. Dorn, R. et al. *Genetics* **133**, 279–290 (1993).
9. Soeller, W. C., Euk Oh, C. & Kornberg, T. B. *Molec. cell. Biol.* **13**, 7961–7970 (1993).
10. Biggin, M. & Tijian, R. *Cell* **53**, 699–711 (1988).
11. Gilmour, D., Thomas, G. & Elgin, S. C. *Science* **245**, 1487–1490 (1989).
12. Soeller, W. C., Poole, S. J. & Kornberg, T. *Genes Dev.* **2**, 68–81 (1988).
13. Thummel, C. *Genes Dev.* **6**, 2635–2645 (1989).
14. Read, D., Nishigaki, T. & Manley, J. L. *Molec. cell. Biol.* **10**, 4334–4344 (1990).
15. Kerrigan, L. A., Croston, G. E., Lira, L. M. & Kadosh, J. T. *J. biol. Chem.* **266**, 574–582 (1991).
16. Harrison, S. H. & Travers, A. A. *EMBO J.* **9**, 207–216 (1990).
17. Chi, H. C., Juminaga, D., Wang, S. Y. & Hui, C. F. *Molec. cell. Biol.* **10**, 451–466 (1991).
18. Lee, J., Kraus, W., Wolfner, M. F. & Lis, J. *Genes Dev.* **6**, 284–295 (1992).
19. Lu, Q., Wallrath, L. L., Granok, H. & Elgin, S. C. *Molec. cell. Biol.* **13**, 2802–2814 (1993).
20. Croston, G. E., Kerrington, L. A., Lira, L. M., Marshak, D. R. & Kadosh, J. T. *Science* **251**, 643–649 (1991).
21. DiBello, P. R., Withers, D. A., Bayer, C. A., Fristrom, J. W. & Guild, G. M. *Genetics* **129**, 385–397 (1991).
22. Godt, D., Couderc, J. L., Cramton, S. E. & Laski, F. A. *Development* **119**, 799–812 (1993).
23. Xue, F. & Cooley, L. *Cell* **72**, 681–693 (1993).
24. Dorn, R., Reuter, G. & Saumweber, H. *Proc. natn. Acad. Sci. U.S.A.* **90**, 11376–11380 (1994).
25. Duncan, I. *Genetics* **102**, 49–70 (1982).
26. Galloni, M., Gyurkovics, H., Schedl, P. & Karch, F. *EMBO J.* **12**, 1087–1097 (1993).
27. Brown, N. & Kafatos, N. *J. molec. Biol.* **203**, 425–437 (1988).

ACKNOWLEDGEMENTS. We thank G. Faustino, A. Apátiné, J. Guiard-Maffia and G. Dongóné for technical assistance, K. Bhat, P. Schedl, P. Spierer and G. Tick for critical reading of the manuscript, F. Cléard for providing developmental northern blots and W. Soeller, T. Kornberg and C. Wu for sharing unpublished results. This work was supported by the Swiss National Fund, the Canton of Geneva and an OTKA grant (920) provided by the Hungarian National Granting Agency. We would especially like to acknowledge and thank the Human Frontier Science Program Organization for a grant which made this collaborative research project possible.

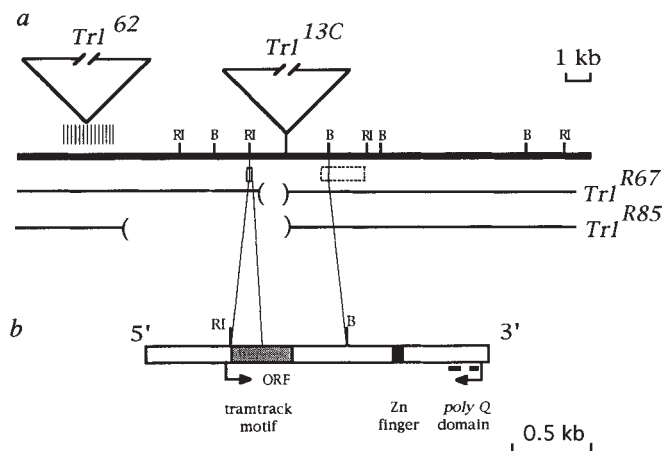


FIG. 3 Molecular map of the *Trithorax*-like locus. *a*, The genomic map is drawn with restriction sites for *EcoRI* (RI) and *BamHI* (B). The positions of the P-elements in *Trl^{13C}* and *Trl⁶²* are represented by triangles. The transcription unit is also indicated below the DNA line. The position of the first exon was determined by sequencing (rectangle); the 3' exon(s) were localized by restriction mapping (rectangle with dotted line). The extent of the deletions in *Trl^{R67}* and *Trl^{R85}* are shown below the genomic map. *b*, The structure of the largest cDNA with the corresponding restriction sites (RI and B) is drawn at a different scale. The open reading frame with the positions of the regions coding for the N-terminal conserved domain, the putative zinc-finger and stretches of polyglutamine residues are also shown (see also ref. 9).

METHODS. The original *Trl^{13C}* allele is from the insertion of a P-element described in ref. 26. Flanking DNA was recovered from a genomic library constructed from *Trl^{13C}/TM2* flies. Wild-type genomic DNA was isolated from the Maniatis library. Using the *EcoRI* fragment spanning the P-element insertion site as a probe, a complex transcription pattern was detected on northern blot of poly(A)⁺ RNA from different embryonic stages with a 2.5-kb major band. The same probe was used to screen cDNA clones from a 4–8-h *Drosophila* embryonic library²⁷. *Trl^{R85}* and *Trl^{R67}* were obtained by mobilization of the P-element in the original *Trl^{13C}* allele. We selected the *Trl^{R85}* and *Trl^{R67}* lines on the basis of their stronger enhancer-of-PEV phenotype when tested with *In(1)W^{m4h}*. The extent of the deletions in *Trl^{R85}* and *Trl^{R67}* was determined by whole-genome Southern analysis.

Crystal structure of photolysed carbonmonoxy-myoglobin

Ilme Schlichting*, Joel Berendzen†, George N. Phillips Jr‡ & Robert M. Sweet§

* Department of Biophysics, Max Planck Institute for Medical Research, Jahnstrasse 29, 69120 Heidelberg, Germany

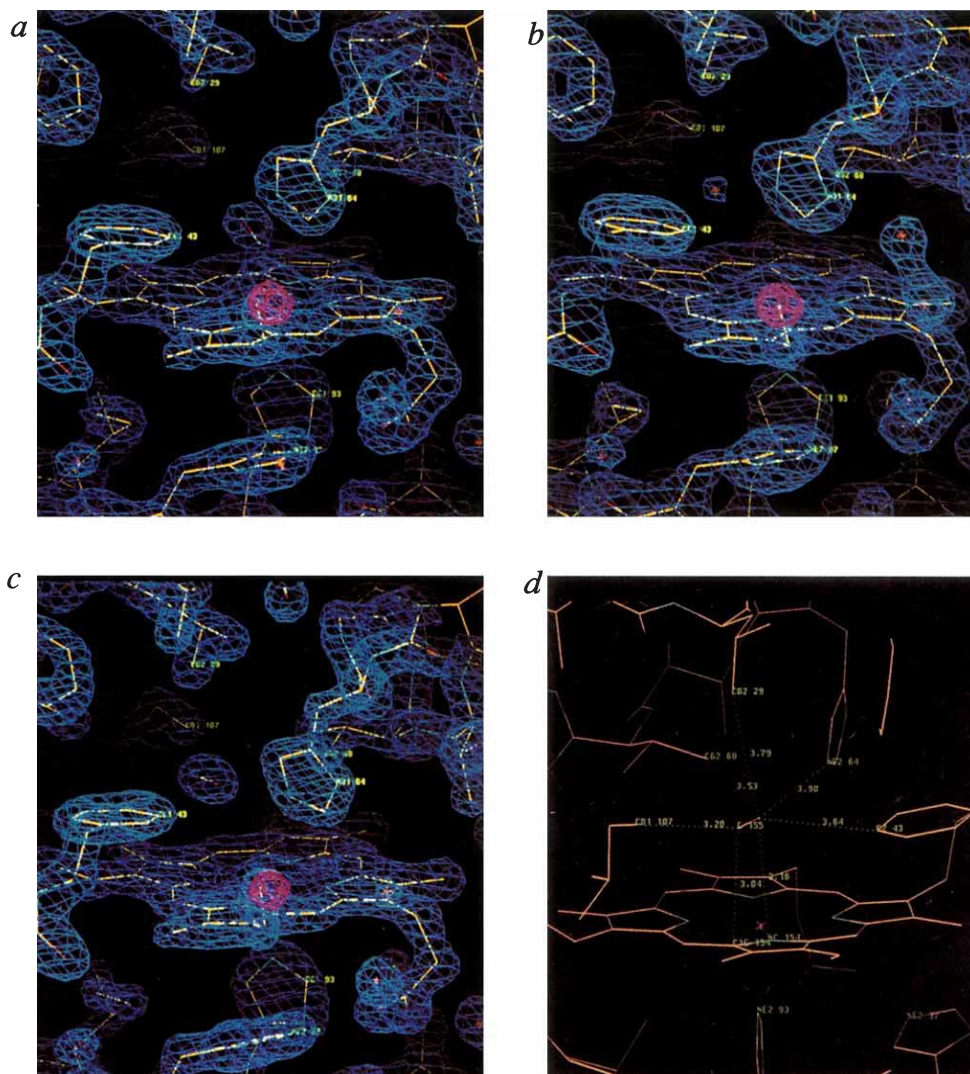
† Department of Biochemistry and Cell Biology and the W.M. Keck Center for Computational Biology, Rice University, Houston, Texas 77251–1892, USA

§ Biology Department, Brookhaven National Laboratory, Upton, New York 11973, USA

MYOGLOBIN is a globular haem protein that reversibly binds ligands such as O₂ and CO. Single photons of visible light can break the covalent bond between CO and the haem iron in carbonmonoxy-myoglobin (MbCO) and thus form an unstable intermediate, Mb^{*}CO, with the CO inside the protein^{1,2}. The ensuing rebinding process has been extensively studied as a model for the interplay of dynamics, structure and function in protein reactions. We have used X-ray crystallography at liquid-helium temperatures to determine the structure of Mb^{*}CO to a resolution of 1.5 Å. The photodissociated CO lies on top of the haem pyrrole ring C. Comparison with the CO-bound and unligated myoglobin structures reveals that on photodissociation of the CO, the haem 'domes', the iron moves partially out of the haem plane, the iron-proximal histidine bond is compressed, the F helix is strained and the distal histidine swings towards the outside of the ligand-binding pocket.

† To whom correspondence should be addressed at Los Alamos National Laboratory, Biophysics Group, MS M715, Los Alamos, New Mexico 87545, USA.

FIG. 1 Active site of myoglobin in three states. $(2F_o - F_c) \exp(i\alpha_c)$ electron density maps are contoured at 16% (blue) and 70% (red) of the maximum electron density. (F_o represents the observed structure factor amplitudes, F_c and α_c are calculated structure factor amplitudes and phases from the refined model.) **a**, MbCO: the haem is flat and the iron is in the plane. We see a single orientation for the CO. The distal histidine is in the pocket with its N ϵ^2 within H-bonding distance of the ligand O. **b**, Unligated Mb: the haem is domed and bent and the iron is 0.32 Å out of the haem plane. Compressing the Fe–N ϵ^2 bond causes the proximal histidine to pivot about a contact point between its C ϵ^2 and the haem and slide the F helix towards the EF corner. The distal histidine is swung slightly outwards and is within H-bonding distance of a water molecule in the pocket. The E helix is bent outwards. **c**, **d**, Mb*CO: the E helix is not bent. There is no density in our Mb*CO structure where the CO had been in MbCO, which constitutes the strongest possible evidence that the crystals were fully photolysed in our experiment. The electron density that corresponds to the photodissociated CO is prolate along what we take to be the mean CO axis. The refined CO has a high occupancy (1.0 C, 0.88 O) and a low *B*-factor (5.9 Å² C, 7.0 Å² O, compared with the overall protein *B*-factor of 6.4 Å²). The photodissociated CO lies atop the haem between the nitrogen and carbon 1 of pyrrole C, close to C ϵ^2 of Phe 43 (CD1), C δ^1 of Ile 107 (G8), C δ^2 of Leu 29 (B10), C γ^1 of Val 68 (E11) and the distal histidine (His 64, E7). Xenon binds at this site³¹ and a water molecule sits here in unligated Mb.



Rebinding of photodissociated CO to myoglobin proceeds by two distinct processes: the CO can either rebind from within an internal cavity called the distal haem pocket, or it can escape and rebind from the solvent. Lowering the temperature slows and separates the processes, until below 180 K the ligand cannot leave the pocket². Although the photodissociated state at ambient and cryogenic temperatures has been studied by many techniques, including visible^{2,3}, near-infrared (NIR)^{3,4}, Raman⁴⁻⁶, infrared (IR)⁷⁻⁹ and extended X-ray absorption fine structure (EXAFS)^{10,11} spectroscopies and by theoretical and computational methods^{9,12-16}, the pieces of the structure–function puzzle have not yet fallen together.

The dynamics of ligand rebinding, together with heat and light transfer in the crystal, places fundamental limitations on crystal thickness, photolysis power, measurement temperature and timescale. Early attempts to obtain the structure of the photodissociated state at room temperature failed because of the difficulty of acquiring X-ray diffraction data on the millisecond timescale¹⁷. Because the presence of appreciable amounts of MbCO in the photolysed crystal would adversely affect the accuracy of the corresponding structure through a reduction in the size of the observed changes, we placed the highest priority on an experimental design that ensures the sample is fully photolysed and remains that way for the duration of the measurement. Our strategy has been to simplify the experiment by conducting it at 20 K, where CO rebinding is negligible over roughly 1 h.

We used optically thin crystals, a large photolysis beam, and continuous photolysis with a photolysis rate coefficient large enough to rephotolyse the crystal quickly in the event of a short-lived temperature fluctuation but small enough to guard against beam heating. Hexagonal (*P*6) crystals of a sperm-whale myoglobin mutant (Asp 122→Asn)¹⁸ grown at pH 9 were used in this study. Hexagonal Mb*CO crystals exhibit CO stretch frequencies and low-temperature rebinding kinetics nearly identical to those in frozen solutions (U. Nienhaus, personal communication).

We collected two X-ray diffraction data sets on Mb*CO crystals at a temperature near 20 K, as well as control data sets from MbCO and unligated myoglobin at 85 K to account for any structural changes associated with the cryoprotectant and low temperatures. Data collection and refinement statistics are shown in Table 1. We refined all three structures in the same way to avoid any spurious differences in structure that might arise from differences in refinement parameters or procedures. Because our MbCO and unligated myoglobin structures, which may be regarded as extremes on either side of the Mb*CO structure, agree with published values, it is unlikely that the accuracy of the photolysed structure is compromised through over-restraint. On photolysis we see major changes in the positions of the CO, distal histidine and haem, as well as smaller changes in the geometry of the proximal histidine, the F helix, Arg 45 and residues in the haem pocket (Figs 1 and 2).

TABLE 1 Data collection and final refinement statistics

	Mb*CO-1	Mb*CO-2	Mb*CO-1 + 2§	MbCO	Unligated Mb
Data collection					
Cell dimensions (Å)	a = b = 90.01, c = 45.20	a = b = 90.42, c = 45.40	a = b = 90.42, c = 45.40	a = b = 90.25, c = 45.31	a = b = 90.62, c = 45.32
Resolution (Å)	1.7	1.5	1.5	1.9	2.0
Observations (no.)	41,952	51,994	93,946	39,961	23,666
Unique reflections, all refl./> 3σ (no.)	19,851/16,288	29,798/24,347	30,947/25,553	14,194/10,341	11,478/9,934
Completeness, all refl./3σ (%)					
∞–6.0 Å	79/79	67/67	86/85	95/90	98/84
6.0–2.9 Å	94/91	87/86	98/97	98/91	99/76
2.9–2.4 Å	91/86	91/87	98/94	88/67	89/31
2.4–2.0 Å	85/75	91/83	98/91	80/49	55/12
2.0–1.9 Å	78/63	89/78	96/85	60/20	
1.9–1.7 Å	66/46	86/69	93/75		
1.7–1.5 Å		78/54	78/54		
R _{sym} (%)†	6.8	4.7	6.3	6.6	12.0
Temperature (K)	≈ 20	≈ 20	≈ 20	85	85
Crystal size (μ)	100 × 100 × 150	150 × 150 × 200		200 × 200 × 400	150 × 150 × 200
Final refinement					
R _{cryst} (%)‡			20.7	21.6	24.3
Resolution (Å)			1.5	1.9	2.1
Waters (no.)			117	97	98
mean B-factor (Å ²)			6.4	7.0	6.3
r.m.s. bond length deviations (Å)			0.010	0.015	0.011
r.m.s. bond angle deviations (°)			1.7	2.5	1.9
r.m.s. coordinate error (est.) (Å)			0.06	0.09	0.13

Crystal preparation. Hexagonal (P6) crystals of a sperm-whale myoglobin mutant (Asp 122 → Asn)¹⁸ were soaked for ~1 h in a thoroughly degassed and N₂-saturated cryoprotectant solution made by addition of 100 mg sucrose, 100 mg glucose and 8 mg ml⁻¹ Na-dithionite to 1 ml of 70% saturated ammonium sulphate with 50 mM Tris-HCl pH 9.0. To generate the MbCO complex, CO was added. Distinct colour changes accompanied the formation of unligated myoglobin (Mb) from met-Mb and MbCO from unligated Mb. The crystals were mounted on a loop²² made of nylon monofilament on which a film was formed of a solution similar to the cryoprotectant except containing 250 mg sucrose ml⁻¹. **Cryogenics.** The crystals were flash-cooled, either *in situ* in a He gas/liquid stream (details to be published elsewhere) in the case of the Mb*CO data, or in liquid nitrogen for the MbCO and unligated Mb data. The liquid helium consumption rate was 4 l h⁻¹ at 20 K. For the Mb*CO data, the temperature was measured by placing a Si diode thermometer in the cryogen stream. **Photolysis.** Photolysis was achieved using white light from a fibre-optic illuminator run continuously at low power for the duration of the data collection. The calculated photolysis rate coefficient was 10^{-1.5 ± 1} s⁻¹.

Diffraction data. X-ray diffraction data for MbCO and unligated Mb were collected with a Siemens multi-wire area detector using focused X-rays from a rotating Cu-anode generator and were reduced with XDS²³. Data sets for two photolysed crystals (Mb*CO-1 and Mb*CO-2, total collection times of 132 and 76 min, respectively) were collected at beamline X12-C of the National Synchrotron Light Source with a MAR Research image plate detector using X-rays of wavelength 0.91 Å and were reduced with DENZO (Z. Otwinowski). **Refinement.** The structure of met-myoglobin in the P6 spacegroup (Protein Data Bank entry 1MBW) was used as a starting model for refinement with X-PLOR²⁴ with the coordinates of the haem and water molecules omitted. After a first round of rigid-body, energy-minimization and B-factor refinements, a 2F_o – F_c electron density map was calculated and inspected²⁵. High electron densities were found for the haem and the CO. After manual rebuilding and insertion of the haem, the CO and water molecules, other rounds of refinement were done that included energy minimization, refinement of individual B-factors and occupancies of the water molecules and CO. All structures were refined with the same set of parameters, with restraints for the haem weakened from the standard X-PLOR values to allow the haem, the ligand, and the proximal histidine to assume positions centred in active-site omit-map densities. All structures were independently refined in two different laboratories and found to agree.

† $R_{\text{sym}} = \sum |I_{hi} - \langle I \rangle_n| / \sum \langle I \rangle_n$; I_{hi} is the scaled intensity of the i th symmetry-related observation of reflection h and $\langle I \rangle_n$ is the mean value.

‡ $R_{\text{cryst}} = \sum |F_{o_i} - F_{c_i}| / \sum F_{o_i}$, where F_{o_i} and F_{c_i} are the observed and calculated structure factor amplitudes for reflection h . We used all reflections and a low-resolution limit of 7 Å for this calculation.

§ The two datasets for Mb*CO were independently refined at first but then combined after we made sure the structures were the same.

|| As estimated by σ_A plot²⁶.

The photodissociated CO sits in the centre of the haem pocket, in agreement with the predictions of computations^{9,13,14,16}. It is not possible to determine from our data which end of the CO points towards the iron. However, it is oriented with respect to the haem, making an angle with the haem plane of $91 \pm 20^\circ$, which agrees with the value of $80 \pm 10^\circ$ determined at 100 ps at 150 K by IR spectroscopy⁹. The apparent discrepancy with the conclusions of ref. 8 that the CO is unoriented in the pocket at 15 K may be due to the low signal-to-noise ratio of that determination.

The geometries of the three structures are compared in Table 2. The distances that we obtain from the iron to the ligand atoms in Mb*CO are 0.7 to 1.6 Å larger than those determined by EXAFS. It is unlikely that the discrepancy can be explained by differences between hydrated films and crystals because the IR C–O stretch frequencies are not significantly different in the two cases¹⁹. Although it is possible that the geometry of some 30% of the photodissociated CO population could change between 4

and 20 K as part of the transition seen in C–O stretch frequencies⁷, most of the population remains the same.

In the photolysed state the distal histidine is positioned towards the outside of the pocket beyond the unligated position, which differs qualitatively from the predictions of molecular dynamics¹⁴. Energy has been stored in the conformation of this residue and its neighbours (such as Arg 45). We speculate that this energy is released during structural fluctuations associated with ligand escape²⁰.

Comparison of the haem geometry among the three states reveals that on photolysis of the CO, the iron moves out of the mean haem plane by 80% of the change between MbCO and unligated myoglobin. The movement of the pyrrole nitrogens out of the haem plane (doming) is complete. These findings are consistent with spectroscopic data on the NIR charge-transfer band III, which is considerably shifted from its unligated position in a direction indicative of smaller iron out-of-plane displacement^{3,4}. There is quantitative agreement with the find-

TABLE 2 Haem-ligand geometries

	Mb*CO structures			MbCO structures				Unligated Mb structures		
	This study	EXAFS ^{11†}	EXAFS ^{10‡}	This study	X-ray ²⁷	Neutron ^{28§}	X-ray ^{29§}	This study	X-ray ²⁷	X-ray ³⁰
CO										
IR angle (°)	91 [¶]			32	19	47/47	40/61			
CO bend angle (°) [#]	111 [¶]	115/140		160	169	135/146	120/141			
Tilt angle (°) ^{**}	39 [¶]			3.9	8.1	13.2	2.7			
Fe-C (Å)	4.14 [¶]	1.97/2.58	2.17/2.35	1.85	1.90	2.21	1.92			
Fe-O (Å)	3.60 [¶]	-/3.42		2.93	3.01	3.08/3.19	2.73/2.93	4.12 ^{††}	3.97 ^{††}	3.78 ^{††}
O-distal His N ^{ε2} (Å)	3.90 [¶]			2.81	3.07	2.60/3.48	3.96/2.66	3.09 ^{††}	2.68 ^{††}	3.05 ^{††}
Haem										
Fe-⟨N _p plane⟩ (Å)	0.19			0.04	0.05	0.06	0.00	0.25	0.32	0.30
Fe-⟨haem plane⟩ (Å)	0.27			0.05	0.10	0.10	0.03	0.32	0.40	0.35
⟨Fe-N _p ⟩ (Å)	1.98	2.03/2.05	1.98/1.96	1.96	1.97	2.02	1.97	1.97	1.97	2.02
⟨N _p -N _p ⟩ (Å)	2.78			2.78	2.79	2.70	2.79	2.76	2.74	2.82
Proximal histidine										
Fe-N ^{ε2} (Å)	2.25	2.22/2.20	2.18/2.19	2.30	2.27	2.26	2.19	2.20	2.25	2.10
Tilt angle (°) [*]	3.3			5.9	4.8	3.5	5.4	8.2	7.8	2.8
dihedral										
N _p A-Fe-N ^{ε2} -C ^{ε2} (°)	12			17	14	-4	-10	13	-1	-2
Distal histidine dihedrals										
χ ₁ (°)	-167			-180	-168	-168	-164	-178	-177	-175
χ ₂ (°)	54			62	66	60	60	66	59	60
χ ₃ (°)	-172			-176	-166	-179	179	-177	-177	-176

Angled brackets indicate a mean value averaged over several atoms.

[†] Pairs of entries refer to the 4 K and 40 K EXAFS structures, which also used different photolysis protocols. The assumption that the C of CO is nearest the iron is in contrast to our assumption, and is not excluded by either set of data.

[‡] Pairs of entries refer to calculated values based on assumptions of 2.12 Å and 2.20 Å, respectively, for the Fe-N^{ε2} bond length in MbCO. The C of CO is nearest to the iron in these calculations.

[§] Pairs of entries refer to alternative CO conformations that were refined in these structures.

^{||} The IR angle lies between the C-O bond and the normal to the mean haem plane.

[¶] Assumes the O of CO is nearest the iron. The other orientation is not excluded by our data. These values have increased uncertainty due to the angular disorder of the CO orientation in the photolysed state.

[#] The CO bend angle is between the iron, the nearer ligand atom and the farther ligand atom.

^{**} The tilt angle lies between the line Fe-ligand and the normal to the mean haem plane.

^{††} These entries refer to distances to the O of a water molecule at the active site.

ings of Dasgupta and Spiro, who estimate the iron out-of-plane relaxation to be 70% in photodissociated carbonmonoxy-haemoglobin within 7 ns at room temperature on the basis of vibrational spectroscopy²¹, and in qualitative agreement with the molecular dynamics calculations of Kuczera *et al.* who find an 80% relaxation within 50 fs at room temperature¹⁵. Any changes in the Fe-pyrrole N distance are too small to be detected by our measurements. EXAFS is equivocal on this point, but Raman studies suggest that the haem core is larger in Mb*CO than in unligated myoglobin and provide additional evidence that the iron is not fully out-of-plane⁶.

We find that the iron-proximal histidine bond shortens on photolysis roughly halfway to its unligated length. This change

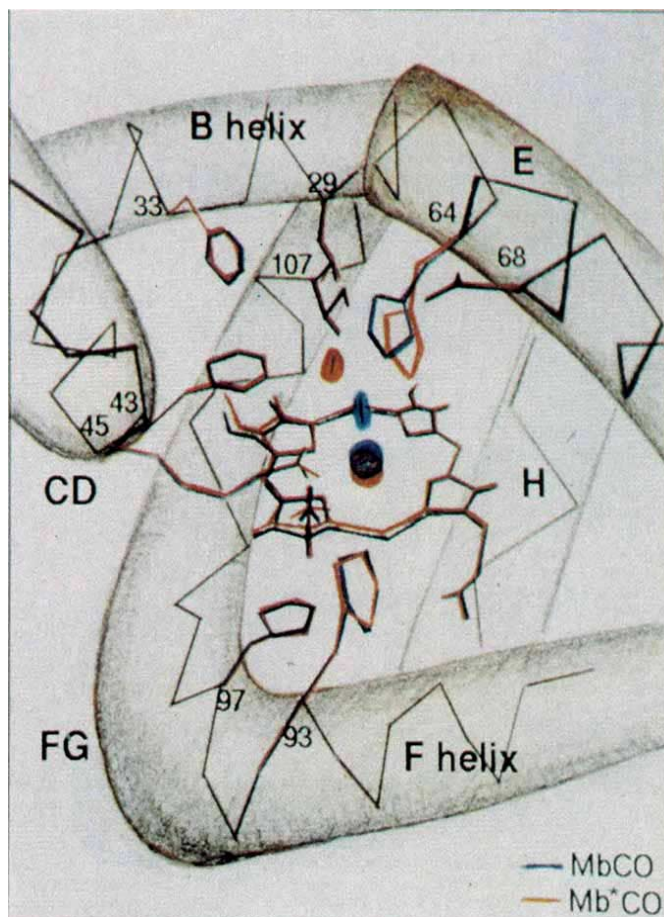


FIG. 2 Differences between MbCO (blue) and Mb*CO (red). The distal histidine in Mb*CO does not lie between the MbCO and unligated positions, but is swung further out of the pocket than in either. The proximal histidine in Mb*CO has moved in a concerted way with the iron towards the unligated position. Neither the Fe-N^{ε2} bond nor the F helix have fully relaxed, and strain appears to be accumulated in the F helix for one or two residues on either side of the proximal histidine. Phe 43 and Leu 29 move slightly on photodissociation of the CO, whereas Val 68 and Ile 107 are unchanged. The carboxyl group of the propionic acid attached to the D pyrrole rotates by some 20°, resulting in a compensating conformational change of its hydrogen-bonding partner Arg 45 (CD3). These groups change more still in going to the unligated state. Arg 45 lies between another Xe binding site and the solvent is suspected of controlling an exit from the pocket⁹.

cannot be directly compared with those seen in the Raman Fe-N² stretch mode on photolysis⁵, as several structural parameters determine its frequency and intensity¹². However, our results can provide a basis for structural interpretation of the complex changes seen in this important spectroscopic marker.

The structure of the photolysed state provides new insight into the dynamics of the protein. Transient spectroscopic studies have shown that Mb*CO relaxes towards the unligated structure on the subnanosecond timescale at physiological temperature. Associated with this relaxation is a dramatic increase in the enthalpic part of the barrier to ligand rebinding from the pocket^{3,4,25}. Comparison of the Mb*CO and unligated conformations allows us to identify the relaxation as an increase in the iron out-of-plane distance, a decrease in the Fe-proximal histidine bond length, a tilting of the proximal histidine and a decompression and motion of the F helix towards its junction with the E helix. These structural changes are functionally important, because an increase in the enthalpic barrier decreases the associated rate coefficient and thus lowers the ligand-binding affinity. The relaxation of the proximal haem environment from Mb*CO to unligated myoglobin is conceptually very similar to the allosteric R→T transition of haemoglobin.

We have used X-ray crystallography at liquid-helium temperatures to obtain snapshots of myoglobin before, during and after its reaction with carbon monoxide. Our results have been compared with information from spectroscopy and computation, and can help to decide the most fruitful approaches to questions about dynamics. Our detailed structural information about the unstable Mb*CO complex may be combined with the results of other approaches to make quantitative structural-kinetic models of the ligand binding process and to build a comprehensive picture of how a simple protein such as myoglobin works. □

Received 7 July; accepted 23 September 1994.

- Antonini, E. & Brunori, M. *Hemoglobin and Myoglobin in Their Reactions with Ligands* (North-Holland, Amsterdam, 1971).
- Austin, R. H., Beeson, K. W., Eisenstein, L., Frauenfelder, H. & Gunsalus, I. C. *Biochemistry* **14**, 5355-5373 (1975).
- Steinbach, P. J. et al. *Biochemistry* **30**, 3988-3999 (1991).
- Ahmed, A. M. et al. *Chem. Phys.* **158**, 329-351 (1991).
- Sassaroli, M., Dasgupta, S. & Rousseau, D. L. *J. biol. Chem.* **261**, 13704-13713 (1986).
- Rousseau, D. L. & Argade, P. V. *Proc. natn. Acad. Sci. U.S.A.* **83**, 1310-1314 (1986).
- Alben, J. O. et al. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3744-3748 (1982).
- Ormos, P. et al. *Proc. natn. Acad. Sci. U.S.A.* **85**, 8492-8496 (1988).
- Rothberg, L. J., Roberson, M. & Jedju, T. M. *Proc. Soc. photo-opt. Instrum. Engrs* **1599**, 309-315 (1991).
- Teng, T.-Y., Huang, H. W. & Olah, G. A. *Biochemistry* **26**, 8066-8072 (1987).
- Powers, L. et al. *Biochemistry* **26**, 4785-4796 (1987).
- Stavrov, S. S. *Biophys. J.* **65**, 1942-1950 (1993).
- Straub, J. E. & Karplus, M. *Chem. Phys.* **158**, 221-248 (1991).
- Gibson, Q. H., Regan, R., Elber, R., Olson, J. S. & Carver, T. E. *J. biol. Chem.* **268**, 17908-17916 (1992).
- Kuczera, K., Lambry, J. C., Martin, J.-L. & Karplus, M. *Proc. natn. Acad. Sci. U.S.A.* **90**, 5805-5807 (1993).
- Case, D. A. & Karplus, M. *J. molec. Biol.* **132**, 343-368 (1979).
- Bartunik, H. D. *Nucl. Instrum. Meth.* **208**, 523-533 (1983).
- Phillips, G. N. Jr, Arduini, R. M., Springer, B. A. & Sligar, S. G. *Proteins Struct. Funct. Genet.* **7**, 358-365 (1990).
- Hong, M. K. thesis, Univ. Illinois at Urbana-Champaign (1989).
- Ringe, D., Petsko, G. A., Kerr, D. E. & Ortiz de Montellano, P. R. *Biochemistry* **23**, 2-4 (1984).
- Dasgupta, S. & Spiro, T. G. *Biochemistry* **25**, 5941-5948 (1986).
- Teng, T.-Y. *J. appl. Crystallogr.* **23**, 387-391 (1990).
- Kabsch, W. *J. appl. Crystallogr.* **26**, 795-800 (1993).
- Brünger, A. T. *X-PLOR (Version 3.1) Manual* (Yale Univ., New Haven, 1993).
- Agmon, N. & Hopfield, J. J. *J. chem. Phys.* **79**, 2042-2053 (1983).
- Read, R. *Acta crystallogr. A* **42**, 140-149 (1986).
- Quillin, M. L., Arduini, R. M., Olson, J. S. & Phillips, G. N. Jr *J. molec. Biol.* **234**, 140-155 (1993).
- Cheng, X. & Schoenborn, B. P. *J. molec. Biol.* **220**, 381-399 (1991).
- Kuriyan, J., Wilz, S., Karplus, M. & Petsko, G. A. *J. molec. Biol.* **192**, 133-154 (1986).
- Phillips, S. E. V. *J. molec. Biol.* **142**, 531-554 (1980).
- Tilton, R. F., Kuntz, I. D. & Petsko, G. A. *Biochemistry* **23**, 2849-2857 (1982).

ACKNOWLEDGEMENTS. We thank G. A. Petsko and D. Ringe for establishing and supporting the collaboration, R. S. Goody and K. C. Holmes for their continued support, K. Chu and M. Quillin for technical assistance and H. Frauenfelder, W. Kabsch, U. Nienhaus, J. Olson, P. Philzackery, B. Schoenborn and M. Soltis for discussions and advice. This work was supported by grants from NATO, the International Human Frontiers Organization, the Alexander von Humboldt Foundation, the NIH, the W.M. Keck Foundation, the Welch Foundation, the US Department of Energy Office of Health & Environmental Research, the National Science Foundation, Brookhaven National Laboratory, Stanford Synchrotron Radiation Laboratory and Los Alamos National Laboratory. Coordinates are being deposited in the Brookhaven Protein Data Bank.

ERRATA

Multiple processing streams in occipitotemporal visual cortex

Edgar A. DeYoe, Daniel J. Felleman,
David C. Van Essen & Evelyn McClendon

Nature **371**, 151-154 (1994)

EDITING of the opening sentence of the first paragraph of this Letter introduced a misleading error: the internal subdivisions given in parentheses should correctly read: "'blobs', 'interblobs' and layer 4B in V1; 'thin', 'thick' and 'interstripes' in V2)". □

The ancient regulatory-protein family of WD-repeat proteins

Eva J. Neer, Carl J. Schmidt, Raman Nambudripad
& Temple F. Smith

Nature **371**, 297-300 (1994)

IN this Progress article the WD-repeating units for the TUP1 protein shown in Table 1 should have been coloured as shown here, and not all in black as published. Blocks are coloured according to their fit to the regular expression in Fig. 1 (black, no misses; blue, one miss). □



CORRECTION

Somatostatin-induced inhibition of neuronal Ca²⁺ current modulated by cGMP-dependent protein kinase

Stephen D. Meriney, D. Bruce Gray
& Guillermo R. Pilar

Nature **369**, 336-339 (1994)

IN this Letter, our citation of the unpublished data of E. Butt and U. Walter was incorrect. We stated that Rp-8-pCPT-cGMPS and Sp-8-Br-cGMPS had been shown to be specific inhibitors and activators of cGMP-dependent protein kinase (cGMP-PK). In fact, Rp-8-pCPT-cGMPS is considered a selective inhibitor of cGMP-PK over cAMP-dependent protein kinase (cAMP-PK). (E. Butt and U. Walter, personal communication). Sp-8-Br-cGMPS is not thought to be selective for cGMP-PK over cAMP-PK as similar analogues do not show significant selectivity in cell-free systems (E.B. and U.W., personal communication). The pharmacological potencies of these compounds have not been determined in intact cells, therefore caution should be used to ensure their appropriate use. □